

GenCore version 5.1.6
Copyright (c) 1991 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 28, 2003, 17:45:31 ; Search time 10243 Seconds
(without alignments)
11562.371 Million cell updates/sec

Title: US-09-981-151a-7

Perfect score: 2895
Sequence: 1 cgcctcggatgaagccccc.....ccgggggtctccctgcga 2895

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2889711 seqs, 2045481386 residues
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.tod.*
36: em.htg.man.*
37: em.htg.vrt.*
38: em.sy.*
39: em.brgo.hum.*
40: em.htgo_mus.*
41: em.htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2895	100.0	2895	6	AX574596	Sequence
2	2675.4	92.4	2997	6	AX574590	Sequence
3	1877.8	64.9	3227	9	HSA315734	Homo sapi
4	1877.8	64.9	3675	6	AX319852	Sequence
5	1877.8	64.9	5610	6	BD170080	Novel ADA
6	1877.8	64.9	8435	6	BD170083	Novel ADA
7	1877.8	64.9	8505	6	BD170084	Novel ADA
8	1874.6	64.8	3675	6	AR229456	Sequence
9	1874.6	64.8	3675	6	BD170564	Novel pro
10	1874.6	64.8	4042	6	AR229457	Sequence
11	1795.6	62.0	2433	6	AX574592	Sequence
12	1695	58.5	2902	6	AX574594	Sequence
13	1634.8	56.5	4234	6	BD170079	Novel ADA
14	1634.8	56.5	4234	9	AB095949	Homo sapi
15	1578.4	54.5	7668	6	BD170085	Novel ADA
16	1381.8	47.7	1668	6	AX481380	Sequence
17	1305.4	45.1	3402	10	BC034739	Mus muscu
18	708	24.3	3866	6	BD170498	Novel egg
19	708.4	24.4	3805	6	AX319854	Sequence
20	706.4	24.4	3443	2	HSA311903	Homo sapi
21	663	22.3	1476	6	AR229455	Sequence
22	477.5	14.4	545	6	BD170082	Novel ADA
23	320.2	11.1	422	6	BD170081	Novel ADA
24	311.6	10.8	104776	9	AC022424	Homo sapi
25	310	10.7	67468	2	AC139448	Homo sapi
26	310	10.7	187084	9	AC010269	Homo sapi
27	296	10.2	82998	2	AC022970	Homo sapi
28	259.2	9.0	192708	9	AC091978	Homo sapi
29	227	7.8	2879	6	BD079092	Novel met
30	227	7.8	2879	6	BD079104	Novel met
31	215.4	7.4	5338	6	BD079094	Novel met
32	215.4	7.4	5338	6	BD079106	Novel met
33	214.8	7.3	5383	6	AX128433	Sequence
34	210.6	7.3	5061	6	E55265	Novel metal
35	208	7.2	3132	6	AX327754	Sequence
36	208	7.2	3312	6	AX319851	Sequence
37	208	7.2	3312	6	E55273	Novel metal
38	208	7.2	3349	6	AX327752	Sequence
39	208	7.2	3400	9	AF163762	Homo sapi
40	206.4	7.1	3377	6	AX467546	Sequence
41	206.2	7.1	3207	6	AX327747	Sequence
42	206.2	7.1	3402	6	AX338539	Sequence
43	206.2	7.1	3403	6	AX327745	Sequence
44	206.2	7.1	3403	6	AX338537	Sequence
45	204.4	7.1	2445	6	AX327757	Sequence

ALIGNMENTS

RESULT 1

AX574596

LOCUS

DEFINITION

AX574596

ACCESSION

AX574596.1

VERSION

GI:27551852

KEYWORDS

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE

AUTHORS

AX574596 Sequence 7 from Patent WO0233087. 2895 bp DNA linear PAT 07-JAN-2003

AX574596 Sequence 7 from Patent WO0233087. 2895 bp DNA linear PAT 07-JAN-2003

AX574596 Sequence 7 from Patent WO0233087. 2895 bp DNA linear PAT 07-JAN-2003

AX574596 Sequence 7 from Patent WO0233087. 2895 bp DNA linear PAT 07-JAN-2003

AX574596 Sequence 7 from Patent WO0233087. 2895 bp DNA linear PAT 07-JAN-2003

AX574596 Sequence 7 from Patent WO0233087. 2895 bp DNA linear PAT 07-JAN-2003

AX574596 Sequence 7 from Patent WO0233087. 2895 bp DNA linear PAT 07-JAN-2003

AX574596 Sequence 7 from Patent WO0233087. 2895 bp DNA linear PAT 07-JAN-2003

AX574596 Sequence 7 from Patent WO0233087. 2895 bp DNA linear PAT 07-JAN-2003

Taupier, R.J., Burgess, C.B., Zerhusen, B.D., Kekuda, R., Spytko, K.A.,
Gangoli, E.A., Fernandes, E.R., and Gorman, L.
Proteins and nucleic acids encoding same
Patent: WO 0233087-A 7 25-APR-2002;
Curagen Corporation (US)

Location/Qualifiers
I .2895
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT 699 a -760 c 787 g 649 t

Query Match 100.0%; Score 2895; DB 6; Length 2895;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2895; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCTCTCGATGAAGCCCGCGCGCGCGATGCGCGGCTTGGCGGCGCTGTGGATGCTG 60
DB 1 CGCTCTCGATGAAGCCCGCGCGCGCGATGCGCGGCTTGGCGGCGCTGTGGATGCTG 60
QY 61 TTGGCGCGAGTGGCGAGCGAGTGGCTCCCGGCGCTCCCAACGAGCGGAAACCGCGG 120
DB 61 TTGGCGCGAGTGGCGAGCGAGTGGCTCCCGGCGCTCCCAACGAGCGGAAACCGCGG 120
QY 121 TCGGAGCAGCTGGAGCGAGTGGCTCCCGGCGCTCCCGGCGCTCCCGGCGCTCCCGG 180
DB 121 TCGGAGCAGCTGGAGCGAGTGGCTCCCGGCGCTCCCGGCGCTCCCGGCGCTCCCGG 180
QY 181 GCGATGTCGCGCTGTCTTTCGCGAGGACACTGCGTGGCGCGCTCCCGGCGCTCCCGG 240
DB 181 GCGATGTCGCGCTGTCTTTCGCGAGGACACTGCGTGGCGCGCTCCCGGCGCTCCCGG 240
QY 241 GCTTGGGAGCCGAGGCTCCCGGCGCTCCCGGCGCTCCCGGCGCTCCCGGCGCTCC 300
DB 241 GCTTGGGAGCCGAGGCTCCCGGCGCTCCCGGCGCTCCCGGCGCTCCCGGCGCTCC 300
QY 301 AAGAGCGCGAGCAGTGAAGTGGAGAACCTGCTCTCAGCAACCTAACACAGAGACA 360
DB 301 AAGAGCGCGAGCAGTGAAGTGGAGAACCTGCTCTCAGCAACCTAACACAGAGACA 360
QY 361 GAAACCCAACTGCTGCTTCTCACTCAATATGACCTGCTCTGCTCAGAGGTTCCAC 420
DB 361 GAAACCCAACTGCTGCTTCTCACTCAATATGACCTGCTCTGCTCAGAGGTTCCAC 420
QY 421 CACAGGGGCGATTACGCTCCCATGAATCATGCCATCAGCGGCGGAGAGACGCTG 480
DB 421 CACAGGGGCGATTACGCTCCCATGAATCATGCCATCAGCGGCGGAGAGACGCTG 480
QY 481 GCGGTGTCGAGGTTGAGTCTTCTTCACTTCCGCTGAAAGGCCCGAGGCGAGCTTCCAC 540
DB 481 GCGGTGTCGAGGTTGAGTCTTCTTCACTTCCGCTGAAAGGCCCGAGGCGAGCTTCCAC 540
QY 541 ATGGATCTGAGGACTTTCCAGCAGCCTAGTGCTCTTGGCTTTATGTGAGAGGTTGGGA 600
DB 541 ATGGATCTGAGGACTTTCCAGCAGCCTAGTGCTCTTGGCTTTATGTGAGAGGTTGGGA 600
QY 601 AAGACAGGCACTAAGTCTGTGAGAGCTTACCGCGAGAGGACTTCTGTTCTATCAAGGC 660
DB 601 AAGACAGGCACTAAGTCTGTGAGAGCTTACCGCGAGAGGACTTCTGTTCTATCAAGGC 660
QY 661 TCTTTTCGATCACACAGAACTCCCATCGATGGAGGGAAGTCTGTGGGGCTCCACT 720
DB 661 TCTTTTCGATCACACAGAACTCCCATCGATGGAGGGAAGTCTGTGGGGCTCCACT 720
QY 721 CGCACTCTGAAGCTCTGCAACAGTCAAGAAATGTCTCCCGGACAGTGTGACTTCGCTGCT 780
DB 721 CGCACTCTGAAGCTCTGCAACAGTCAAGAAATGTCTCCCGGACAGTGTGACTTCGCTGCT 780
QY 781 GCTCAGTGTGCGGACACACAGACAGAGATTACAGGCGCGCACTACAGTGGAGGCT 840
DB 781 GCTCAGTGTGCGGACACACAGACAGAGATTACAGGCGCGCACTACAGTGGAGGCT 840

QY 841 TACACTCAAGTAGAGCCGACTTATGCAAACTCTACTGTATCGCAGAGGATTTGATTC 900
DB 841 TACACTCAAGTAGAGCCGACTTATGCAAACTCTACTGTATCGCAGAGGATTTGATTC 900
QY 901 TTTCTTTCTTTGTTCAATAAAGTCAAAAGTGGGACTCCATGCTCGAGGATAGCCGTAAT 960
DB 901 TTTCTTTCTTTGTTCAATAAAGTCAAAAGTGGGACTCCATGCTCGAGGATAGCCGTAAT 960
QY 961 GTTTGTATAGTGGGATATGAGCTCAGTGTGTTCCAGTCTGCCACATGCCCCAG 1020
DB 961 GTTTGTATAGTGGGATATGAGCTCAGTGTGTTCCAGTCTGCCACATGCCCCAG 1020
QY 1021 CTTCCCAAGGAAGACCTCTTTCATCTTGGCCAGATGAGTATAAGTCTTTGCTTACGGCATAG 1080
DB 1021 CTTCCCAAGGAAGACCTCTTTCATCTTGGCCAGATGAGTATAAGTCTTTGCTTACGGCATAG 1080
QY 1081 CGCTCTCTCTTGTGAGGTCCTCATAGAAATGAAGAACTGAACTGGAGACTTGGTGGTGC 1140
DB 1081 CGCTCTCTCTTGTGAGGTCCTCATAGAAATGAAGAACTGAACTGGAGACTTGGTGGTGC 1140
QY 1141 GACAAAGATGATGCAAAACCAATGGCCATCAAAATATCAACCACTPACGTCCTCAGCAT 1200
DB 1141 GACAAAGATGATGCAAAACCAATGGCCATCAAAATATCAACCACTPACGTCCTCAGCAT 1200
QY 1201 CTCAATGTGATCTCTTTTATTCAAAGATGGATGATGGGAAAGATGGGACTCTGAT 1260
DB 1201 CTCAATGTGATCTCTTTTATTCAAAGATGGATGATGGGAAAGATGGGACTCTGAT 1260
QY 1261 GACACGCGCTCTTACGATGCTGTGATATATGTTCTTGGAAAGATGAGCCCTGTGAC 1320
DB 1261 GACACGCGCTCTTACGATGCTGTGATATATGTTCTTGGAAAGATGAGCCCTGTGAC 1320
QY 1321 ACTTTGGGATTTGCAACCAATGAGTGGATGATGATGATGATGATGATGATGATGAT 1380
DB 1321 ACTTTGGGATTTGCAACCAATGAGTGGATGATGATGATGATGATGATGATGATGAT 1380
QY 1381 GAGATGACAGCTCTTGGGACTGGCTTACGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
DB 1381 GAGATGACAGCTCTTGGGACTGGCTTACGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
QY 1441 ATGATGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1500
DB 1441 ATGATGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1500
QY 1501 ACATTTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1560
DB 1501 ACATTTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1560
QY 1561 AAATTTCTAAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1620
DB 1561 AAATTTCTAAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1620
QY 1621 TACAAGTATCTGAGAAATTCGAGGAGAAATATATGATGAGGAGGAGGAGGAGGAGGAG 1680
DB 1621 TACAAGTATCTGAGAAATTCGAGGAGAAATATATGATGAGGAGGAGGAGGAGGAGGAG 1680
QY 1681 CAGTTTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1740
DB 1681 CAGTTTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1740
QY 1741 CTGTGGTGCATCTGATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAAT 1800
DB 1741 CTGTGGTGCATCTGATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAAT 1800
QY 1801 ACAATTTGTGGGCAATGACATGTGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1860
DB 1801 ACAATTTGTGGGCAATGACATGTGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1860
QY 1861 GGGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAG 1920
DB 1861 GGGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAG 1920
QY 1921 ACCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1980

Db	1921	ACCTGCGAGGGGGAGTATCTCATAGAGTGCCTCTGCACCAACCCCAAGCCATGCAT	1980
Qy	1981	GGAGGGAAGTTCTGTGAGGGCTCCACATCGCAGCTCTGAAGCTCTGCAACAGTCAGAAATGT	2040
Db	1981	GGAGGGAAGTTCTGTGAGGGCTCCACATCGCAGCTCTGAAGCTCTGCAACAGTCAGAAATGT	2040
Qy	2041	CCCCGGGACAGTGTGTGACTTCGCGTCTGCTCAGTGTGCCGAGCAACAACAGACAGACATTC	2100
Db	2041	CCCCGGGACAGTGTGTGACTTCGCGTCTGCTCAGTGTGCCGAGCAACAACAGACAGATTC	2100
Qy	2101	AGAGGGCGGCACTCAAGTGGGAAGCTTACACTCAAGTAGAAGATCAGGACTTATGCAAA	2160
Db	2101	AGAGGGCGGCACTCAAGTGGGAAGCTTACACTCAAGTAGAAGATCAGGACTTATGCAAA	2160
Qy	2161	CTCTACTGTATCGGACGAAGGATTGTATTCTTCTTTTCTTGTCTCAATAAAGTCAAAAGT	2220
Db	2161	CTCTACTGTATCGGACGAAGGATTGTATTCTTCTTTTCTTGTCTCAATAAAGTCAAAAGT	2220
Qy	2221	GCGACTCCATGCTCGGAGGATGACCGTATATGTTTGTATATGATGGGATATGTATGAGAGTT	2280
Db	2221	GCGACTCCATGCTCGGAGGATGACCGTATATGTTTGTATATGATGGGATATGTATGAGAGTT	2280
Qy	2281	GGATGTGACAAATGTCCTTCGATCTGATGTGTTCGACAGCGTCTGTGGGTGTGTACGGG	2340
Db	2281	GGATGTGACAAATGTCCTTCGATCTGATGTGTTCGACAGCGTCTGTGGGTGTGTACGGG	2340
Qy	2341	AATAACTCAGGCTCGCAGATTCACAGAGGGTCTCTACACCAAGCACACCAACACACACAG	2400
Db	2341	AATAACTCAGGCTCGCAGATTCACAGAGGGTCTCTACACCAAGCACACCAACACACACAG	2400
Qy	2401	TATTATCACATGGTCACCAATTCCTTCTGGAGCCCGGAGTATCCGCACTCTATGAAATGAAC	2460
Db	2401	TATTATCACATGGTCACCAATTCCTTCTGGAGCCCGGAGTATCCGCACTCTATGAAATGAAC	2460
Qy	2461	GTCTCTACTCTCATATTTCTGTGCGCAATGCCCTCAGAAGGTACTACTCGAATGGGCAC	2520
Db	2461	GTCTCTACTCTCATATTTCTGTGCGCAATGCCCTCAGAAGGTACTACTCGAATGGGCAC	2520
Qy	2521	TGGACGCTGAGCTGGCCCGGCGCGGTACAAATTTTGGGGCACTACTTTCGACTCAGACGG	2580
Db	2521	TGGACGCTGAGCTGGCCCGGCGCGGTACAAATTTTGGGGCACTACTTTCGACTCAGACGG	2580
Qy	2581	TCCTATATATGAGCCCGAGAACTTAATCGCTACTTGGACCAACCAACAGACACTGATTTGTG	2640
Db	2581	TCCTATATATGAGCCCGAGAACTTAATCGCTACTTGGACCAACCAACAGACACTGATTTGTG	2640
Qy	2641	GAGCTGCTGTTTCTAGGAAAGGAAACCCGGGTGTCCCTGGGAAATCTCATGCCCTGCGTTG	2700
Db	2641	GAGCTGCTGTTTCTAGGAAAGGAAACCCGGGTGTCCCTGGGAAATCTCATGCCCTGCGTTG	2700
Qy	2701	GGGACCTGATAGAGGCCCCCTGCCACCGCAGTACACTTGGGCGCATTCGCGCTCTGAG	2760
Db	2701	GGGACCTGATAGAGGCCCCCTGCCACCGCAGTACACTTGGGCGCATTCGCGCTCTGAG	2760
Qy	2761	TGCTTCGCTGCTCGGAGAGGGGTAGTCCCTCTACAGTGCCTCTCTCGGAGGACGATGT	2820
Db	2761	TGCTTCGCTGCTCGGAGAGGGGTAGTCCCTCTCTCGAGTGCCTCTCTCGGAGGACGATGT	2820
Qy	2821	CAGCCTTCAGGCACTCGTATATGCACTGGCCTTTCTTGAACTCTATGAGCAGCCGG	2880
Db	2821	CAGCCTTCAGGCACTCGTATATGCACTGGCCTTTCTTGAACTCTATGAGCAGCCGG	2880
Qy	2881	GGCTTCTCCCTGCCA 2895	
Db	2881	GGCTTCTCCCTGCCA 2895	

RESULT 2	AX574590	AX574590	2997 bp	DNA	linear	PAT 07-JAN-2003
LOCUS						
DEFINITION			Sequence 1	from Patent WO0233087.		
ACCESSION			AX574590			

VERSION	AX574590.1	GI:27551849
KEYWORDS	Homo sapiens (human)	
SOURCE	Homo sapiens	
ORGANISM	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.	
REFERENCE	Edinger,S., Gerlach,V., Macdougall,J.R., Malyankar,U.M., Smithson,G., Millet,I., Peyman,J.A., Stone,D.J., Gunther,E., Ellerman,K., Shmets,R.A., Padigaru,M., Guo,X., Patirajan,M., Taupier,R.J., Burgess,C.E., Zerhusen,B.D., Kekuda,R.S., Spyrek,K.A., Gangoli,E.A., Fernandes,E.R. and Gorman,L. Gangoli,E.A., Fernandes,E.R. and Gorman,L. Patent: WO 023087-A 1 25-APR-2002; Curagen Corporation (US).	
TITLE	Location/Qualifiers	
FEATURES	source 1..2997 /organism="Homo sapiens" /mol_type="Genomic DNA" /db_xref="taxon:9606"	
BASE COUNT	728 a 809 g 673 t 4 others	
ORIGIN		
Query Match	92.4%; Score 2675.4; DB 6; Length 2997;	
Best Local Similarity	94.9%; Pred. No. 0;	
Matches 2866;	Conservative 0; Mismatches 5; Indels 150; Gaps 4;	
Qy	1 GCCTCTGTGAATGAAGCCGCCCGCGCGTGGCGGGCTTCCTCACACAGCGGGAAACCGCGG 60	
Db	1 GCCTCTGTGAATGAAGCCGCCCGCGCGTGGCGGGCTTCCTCACACAGCGGGAAACCGCGG 60	
Qy	61 TTGGCCGAGGTGGCCGACGAGTGAGTCTCCCGGGCGTTCCTCACACAGCGGGAAACCGCGG 120	
Db	61 TTGGCCGAGGTGGCCGACGAGTGAGTCTCCCGGGCGTTCCTCACACAGCGGGAAACCGCGG 120	
Qy	121 TCCGGAACACTTGGAGCGAGTCTCCCGGGCGTTCCTCACACAGCGGGAAACCGCGG 180	
Db	121 TCCGGAACACTTGGAGCGAGTCTCCCGGGCGTTCCTCACACAGCGGGAAACCGCGG 180	
Qy	181 CGCATGTGCCTGTTTTCCTCGACGACCTTGGCGGCATGGGACCCGACGCGCAGC 240	
Db	181 CGCATGTGCCTGTTTTCCTCGACGACCTTGGCGGCATGGGACCCGACGCGCAGC 240	
Qy	241 CCTTGGGAGCCCGAGGCTCCCGGTCCTCTCCACCCTCGGACCGCGCGGCTGGATGGA 300	
Db	241 CCTTGGGAGCCCGAGGCTCCCGGTCCTCTCCACCCTCGGACCGCGCGGCTGGATGGA 300	
Qy	301 AAAGGGCGGGACATGATGAAGCTTGAACAACCATGTTCTCAGCAAACTAACACAGAACA 360	
Db	301 AAAGGGCGGGACATGATGAAGCTTGAACAACCATGTTCTCAGCAAACTAACACAGAACA 360	
Qy	361 GAATAACAAACACTGCATGATGTTCTACTCAATATGACCTGGTCTCTGCTTACGAGGTTGAC 420	
Db	361 GAATAACAAACACTGCATGATGTTCTACTCAATATGACCTGGTCTCTGCTTACGAGGTTGAC 420	
Qy	421 CACAGGGCGGATTACGTGTCCTGAAATCATGACCATCAGCGCGGAGGAAGACGATG 480	
Db	421 CACAGGGCGGATTACGTGTCCTGAAATCATGACCATCAGCGCGGAGGAAGACGATG 480	
Qy	481 GCCGTGTCGAGGTTGAGTCTCTTCCACTTCCGCTGAAAAGGCCCCACAGGCACGACTTCAC 540	
Db	481 GCCGTGTCGAGGTTGAGTCTCTTCCACTTCCGCTGAAAAGGCCCCACAGGCACGACTTCAC 540	
Qy	541 ATGGATCTGAGGACTTCCAGCAGCCTTAGTGGCTCTCTGGCTTTATTGTGCAGAGTTGGGA 600	
Db	541 ATGGATCTGAGGACTTCCAGCAGCCTTAGTGGCTCTCTGGCTTTATTGTGCAGAGTTGGGA 600	
Qy	601 AAGACAGGCACTAAGTCTGTGACAGACTTTATCCGCCACAGGACTTCTGTTTTCTATCAGGC 660	
Db	601 AAGACAGGCACTAAGTCTGTGACAGACTTTATCCGCCACAGGACTTCTGTTTTCTATCAGGC 660	
Qy	661 TCTTTCCGATCACACAGAAACTCGCATCGATGGAGGAAGTTCTGTGAGGSGCTCCACT 720	

Db 661 TCTTTGCGATCACAAGAACTGCCATGCGATGGAGGGAAGTCTGTGAGGGCTCCACT 720
 Qy 721 CGCACTCTGAGACTCTGCAACAGTCAAGAAATGTCCTCCGGGACAGTGTGACTTCGGTCT 780
 Db 721 CGCACTCTGAGACTCTGCAACAGTCAAGAAATGTCCTCCGGGACAGTGTGACTTCGGTCT 780
 Qy 781 GCTGAGTGTGCGGAGCAACAACAGCAAGTTCAGAGGGGGCGCACTACAAGTGGAGGCT 840
 Db 781 GCTGAGTGTGCGGAGCAACAACAGCAAGTTCAGAGGGGGCGCACTACAAGTGGAGGCT 840
 Qy 841 TACACTCAGATGAGAGCCGACTTATGCAAACTCTACTGTATGCGAGAGGATTTGATTC 900
 Db 841 TACACTCAGATGAGAGGNGACTTATGCAAACTCTACTGTATGCGAGAGGATTTGATTC 900
 Qy 901 TCTCTTTCTTCTGCAAAATGATCAAGATGGGACTCAATGCTCGGAGGATAGCGGTAAT 960
 Db 901 TCTCTTTCTTCTGCAAAATGATCAAGATGGGACTCAATGCTCGGAGGATAGCGGTAAT 960
 Qy 961 GTTTGTATAGATGGGATATGTGAGCTCAGTGTGTGTGTCACATCTGCGCAKATGCCCGAG 1020
 Db 961 GTTTGTATAGATGGGATATGTGAGCTCAGTGTGTGTGTCACATCTGCGCAKATGCCCGAG 1020
 Qy 1021 CTTCCGAGAGAGCTCTCTGCTGCGAGATGAGTATAGTCTTGTCTTACGGCATAG 1080
 Db 1021 CTTCCGAGAGAGCTCTCTGCTGCGAGATGAGTATAGTCTTGTCTTACGGCATAG 1080
 Qy 1081 CGCTCTCTTCTGAGGTCCTCAATTAAGATGAGAGACTGAACTGAGAGCTGTGTGTGTGTC 1140
 Db 1081 CGCTCTCTTCTGAGGTCCTCAATTAAGATGAGAGACTGAACTGAGAGCTGTGTGTGTGTC 1140
 Qy 1141 GACAAAAGATGATGCAAAACCATGCGCATGAAAATATCAACCTCACTGCTCAAGTA 1200
 Db 1141 GACAAAAGATGATGCAAAACCATGCGCATGAAAATATCAACCTCACTGCTCAAGTA 1200
 Qy 1201 CTCACAGATGATGCTGCTTTATGAGAA----- 1229
 Db 1201 CTCACAGATGATGCTGCTTTATGAGAA----- 1229
 Qy 1230 ----- 1229
 Db 1261 ATTGTAGGTCGATCTCTTAGAAGATGAACAGCCAGGACTGGTGAATAGTCAACAGCA 1320
 Qy 1230 ----- TGGATTGATGGGGAAGATCGGACT 1254
 Db 1321 GACCACACTTAAGTACCTCTCGCAGTGGCAGTCTGGATTTGATGGGAAAAGTGGGACT 1380
 Qy 1255 CGTCATGACCGACGCTCTTACTGACTGCTGTGATATATGTTCTCGGAGATGACGCC 1314
 Db 1381 CCGTATGACCGACGCTCTTACTGACTGCTGTGATATATGTTCTCGGAGATGACGCC 1440
 Qy 1315 TGTGACACTTTGGGATTTGCACCACTGAAGTGAATGTAGTAATATCCGAGCTGCACG 1374
 Db 1441 TGTGACACTTTGGGATTTGCACCACTGAAGTGAATGTAGTAATATCCGAGCTGCACG 1500
 Qy 1375 ATTAATGAAGATACAGGCTTTGGAATGGCTTACCAATTCGCCATGAGTCTGGACACAC 1434
 Db 1501 ATTAATGAAGATACAGGCTTTGGAATGGCTTACCAATTCGCCATGAGTCTGGACACAC 1560
 Qy 1435 TTTGGCATGATTCATGATGAGAGGGAACATGTGTAAGAAATCCGAGGCGCAACTGTC 1494
 Db 1561 TTTGGCATGATTCATGATGAGAGGGAACATGTGTAAGAAATCCGAGGCGCAACTGTC 1620
 Qy 1495 TCCCTACATTTGGCAGGACCAATGAGATCTTCTCTGCTGCTCACTCCGAGGCGCCAGTAT 1554
 Db 1621 TCCCTACATTTGGCAGGACCAATGAGATCTTCTCTGCTGCTCACTCCGAGGCGCCAGTAT 1680
 Qy 1555 CTACACAAATTTCTAAGCAACCTCAAGACTATCTGCTGCTGATGACGCGCAAGCTGTG 1614
 Db 1681 CTACACAAATTTCTAAGCAACCTCAAGACTATCTGCTGCTGATGACGCGCAAGCTGTG 1740
 Qy 1615 AAGGAATACAGATATCTCTGAGAAATTCGAGGAGGAATATATATGATGCAACACACATGTC 1674
 Db 1741 AAGGAATACAGATATCTCTGAGAAATTCGAGGAGGAATATATATGATGCAACACACATGTC 1800

Qy 1675 AAGTGGCACTTGGAGAGAAAGCCAAAGCTCTGCACTGCTGGACTTTAAAGAGCAATCTGT 1734
 Db 1801 AAGTGGCACTTGGAGAGAAAGCCAAAGCTCTGCACTGCTGGACTTTAAAGAGCAATCTGT 1860
 Qy 1735 AAAGCCCTGTGTGCTCATCTGATTTGGAAGAAATGTGAGACTAAATTTATGCCAGCAGCA 1794
 Db 1861 AAAGCCCTGTGTGCTCATCTGATTTGGAAGAAATGTGAGACTAAATTTATGCCAGCAGCA 1920
 Qy 1795 GAAGGACCAAAATTTGTGGGCATGACATGTGTCGCCGGGAGGACAGTGTGTGAATATATGCT 1854
 Db 1921 GAAGGACCAAAATTTGTGGGCATGACATGTGTCGCCGGGAGGACAGTGTGTGAATATATGCT 1980
 Qy 1855 GATGAAGGCCCAAGACCCCAACCATGCGCACTGCTGTGCGGACTGTCTTCTTGTGTCCTCATG 1914
 Db 1981 GATGAAGGCCCAAGACCCCAACCATGCGCACTGCTGTGCGGACTGTCTTCTTGTGTCCTCATG 2040
 Qy 1915 TCCAGGACTCTGGAGGGGGAGTATCTCATAGGAGTGGCTCTGCGACCAACCCCAAGCACA 1974
 Db 2041 TCCAGGACTCTGGAGGGGGAGTATCTCATAGGAGTGGCTCTGCGACCAACCCCAATGCA 2100
 Qy 1975 TCGCATGAGGAGAAATTTGTGTGAGGGGTCCTACTGCGACTGTGAAGCTGTGCAACAGTCAG 2034
 Db 2101 TCGCATGAGGAGAAATTTGTGTGAGGGGTCCTACTGCGACTGTGAAGCTGTGCAACAGTCAG 2160
 Qy 2035 AAATGTCCCGGGAGCAAGTGTGACTTTCGCTGCTCTCACTGCGGAGGAGCAACAGCAGA 2094
 Db 2161 AAATGTCCCGGGAGCAAGTGTGACTTTCGCTGCTCTCACTGCGGAGGAGCAACAGCAGA 2220
 Qy 2095 CGATTCAGAGGGCGGCACTACAGTGGAAAGCTTTACACTCAAGTAGAGATCAGGACTTA 2154
 Db 2221 CGATTCAGAGGGCGGCACTACAGTGGAAAGCTTTACAGTGGAAAGCTTTACAGTGGAAAGCT 2262
 Qy 2155 TGCATACTCTGATTCGCAAGAGAGATTTGATTTCTTCTTCTTGTGCAAAATAAGTCT 2214
 Db 2263 TGCATACTCTGATTCGCAAGAGAGATTTGATTTCTTCTTCTTGTGCAAAATAAGTCT 2322
 Qy 2215 AAAGATGGGACTCTCACTGCTCGGAGGATAGCCGTAATGTTGTATAGATGGGATATGTGAG 2274
 Db 2323 AAAGATGGGACTCTCACTGCTCGGAGGATAGCCGTAATGTTGTATAGATGGGATATGTGAG 2382
 Qy 2275 AAGATGGGATGTGCAAAATGCTCTTGGATCTGATCTGTTGAAGAGCTCTGTGGGGTGTGT 2334
 Db 2383 ---NTTGGATGTGCAAAATGCTCTTGGATCTGATCTGTTGAAGAGCTCTGTGGGGTGTGT 2439
 Qy 2335 AACGGGAATPACTCAGCTCGACGATTCACAGGGGCTCTACACCAAGCAGCACAACCC 2394
 Db 2440 AACGGGAATPACTCAGCTCGACGATTCACAGGGGCTCTACACCAAGCAGCACAACCC 2499
 Qy 2395 AACGATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2454
 Db 2500 AAC---ATTATCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2556
 Qy 2455 AAGGAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2514
 Db 2557 AAGGAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2616
 Qy 2515 GGGCAGCTGAGGCTGAGCTGGCGCGCGGCTGCAAAATTTTGGGCACTACTTTGACATAC 2574
 Db 2617 GGGCAGCTGAGGCTGAGCTGGCGCGCGGCTGCAAAATTTTGGGCACTACTTTGACATAC 2676
 Qy 2575 AGAGGCTCTTAATGAGCGCGGAGAACTTAATGCTTACTGAGCAACCAAGAGAGAGCTG 2634
 Db 2677 AGAGGCTCTTAATGAGCGCGGAGAACTTAATGCTTACTGAGCAACCAAGAGAGAGCTG 2736
 Qy 2635 ATTTGGAGCTGTGTTTTCAGGGAGGAAACCGGGGTTGTTGCTTGGGAATPACTCATGCTCT 2694
 Db 2737 ATTTGGAGCTGTGTTTTCAGGGAGGAAACCGGGGTTGTTGCTTGGGAATPACTCATGCTCT 2796
 Qy 2695 CGCTTGGGAGCGAGAGGAGCAGCCCTGCGCAGGCTACACTGAGGCGCTGTTGAGGCGC 2754
 Db 2797 CGCTTGGGAGCGAGAGGAGCAGCCCTGCGCAGGCTACACTGAGGCGCTGTTGAGGCGC 2856


```

QY 2755 TCTGAGTGTCTCGTGTCTGCGAGGGGAGGTAGCTGCCTTCCAGTGTCTGCTCTCTGAGGCA 2814
Db 2857 TCTGAGTGTCTCGTGTCTGCGAGGGGAGGTAGCTGCCTTCCAGTGTCTGCTCTCTGAGGCA 2916
QY 2815 GCATGTCAACCTTCAGCCATGCGTCAATGACATGACATGCGCTTCTTGAATCCTAATGAGCA 2874
Db 2917 GCATGTCAACCTTCAGCCATGCGTCAATGACATGACATGCGCTTCTTGAATCCTAATGAGCA 2976
QY 2875 GCCCGGGGCTTCTCCCTGCA 2895
Db 2977 GCCCGGGGCTTCTCCCTGCA 2997

RESULT 3
HSA315734 3227 bp mRNA linear PRI 01-MAR-2002
LOCUS HSA315734
DEFINITION Homo sapiens mRNA for metalloprotease disintegrin 16 (ADAMTS16
gene).
ACCESSION AJ315734
VERSION AJ315734.1 GI:19171177
KEYWORDS ADAMTS16 gene; disintegrin; metalloprotease; thrombospondin.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1
AUTHORS Cal.S.; Obaya.A.J.; Llamazares.M.; Garabaya.C.; Quesada.V. and
Lopez-otin.C.
TITLE Cloning, expression analysis, and structural characterization of
seven novel human ADAMTSs, a family of metalloproteinases with
disintegrin and thrombospondin-1 domains
JOURNAL Gene 283 (1-2), 49-62 (2002)
MEDLINE 21856482
PUBMED 11867212
REFERENCE 2 (bases 1 to 3227)
AUTHORS Cal.S.
TITLE Direct Submission
JOURNAL Submitted (26-JUN-2001) Cal S., Biochemistry and Molecular Biology,
University of Oviedo, Campus del Cristo, Asturias. 33006, SPAIN
FEATURES
source
1..3227
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..3227
/gene="ADAMTS16"
1..3219
/gene="ADAMTS16"
/codon_start=1
/product="metalloprotease disintegrin 16 with
thrombospondin type 1 motif"
protein_id="CA38605.1"
/db_xref="GI:19171177"
/translation="MTVSRVSWGLALMLLAQVAPACAPACAPAAAPGPSVPR
PPFAPRFGMEKGEYDVSVEVDHEDYUSHEIMHQBPAVAVSEVSHLRLK
GPRHDPMDLRTSSGLVYVGLVGTGAGTASVOTLPDPFCVQGISRSHNSVA
LSTCGELSMRIETEDAFPLPGLPSHLWGLRAAGSSPSHLVYKSTBEPHAPGSE
LVSTCMELAPQVLRHSSDLRLGPKQHPGRKKYHFPQPRDLIIPDEYKSLR
HGSLLSRHNEELVYLVVDKRMQNHENITTVLLMLWYALFDGLDLC
INTAIVGLLLEDEQGLVISHADHTLSSFCQNSGLMGKRGTRHDAHLTGLDIC
SWNRFCDTGFAPISGCKSKIRSCITNEDTGLAFTAHSESHGFRNHDGEMNC
KSGEINMPTLAGRNVFSFSCRLQRLFLSTALCLADQPKPVKRIPEKLP
GELYDANTCKMQGEMAKLDFKQALCMKCRGRKCEKFPNAPGAPGCHD
MWKRGQCVKFGDEGPKFTHGSDMSWSPCSTCGGVSHRSRLCTNPFAPSGGR
CEUSTRLKNCQKCPDSVFRFAACAEHNSRRFRGHRTKMPFTIQEDQDCLAY
CIAGDFDFEELSNKYDQSTFCSDSRVCTIDGICERVQCDNLGSDAVEDVSGNG
NNSACTIHRGTQKHHTQYTHRWTFIPSGARSRIYENNVSTSLVRYNRLRYLNLN
MGVTDVMPRTYKSGTDFRSTINENPENLITAGPTNETLIVLLPQGRNVGAMYS
PMTGLVFKQPPAPQSTIWAYVSECSVCGGQVTRGCTROLKTFQNMVSPKPIR
PAPSPKQNSQSCFPAWSGAPNAECSHTCGKRRRRAVACKSNFNSARAQLLPDAVC
TSSPFRHMRHQLQKCHPKKQMLVMSQVGLVSNERG"
776 a 871 c 898 g 662 t
BASE COUNT

```

ORIGIN

```

Query Match 64.9%; Score 1977.8; DB 9; Length 3227;
Best Loca Similarity 84.4%; Pred No. 0;
Matches 2289; Conservative 0; Mismatches 177; Indels 245; Gaps 6;

QY 202 CGAGCACACTCGTGTGCGCATGGAGCCCGACAGCGGAGCGCTGGAGGCCGAGCGTCCC 261
Db 69 GCAGGCACTCGTGTGCGCATGGAGCCCGACAGCGGAGCGCTGGAGGCCGAGCGTCCC 128
QY 262 GCGTCTCTCTCCACCGCGGAGCGCGGGCTGATGGAAGGCGGACATGATGMA 321
Db 129 GCGTCTCTCTCCACCGCGGAGCGCGGGCTGATGGAAGGCGG----- 175
QY 322 GCTGGAAACCATGTTCTTCAGCAAACTAACACAGAAACCAAACTGCTCATGTT 381
Db 176 ----- 175
QY 382 CTCACCTCATATGACTGTCTCTGCTACGAGGTTGACCAACGCGGCGATTTAGGTGTC 441
Db 176 -----AATATGACTGTCTCTGCTACGAGGTTGACCAACGCGGCGATTTAGGTGTC 228
QY 442 CATGAATCATGCACTCATGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 501
Db 229 CATGAATCATGCACTCATGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 288
QY 502 CTTCACTCTGCGCTGAAGGCGCCGACGACATTCACATGATCTGAGGACTTCCAGC 561
Db 289 CTTCACTCTGCGCTGAAGGCGCCGACGACATTCACATGATCTGAGGACTTCCAGC 348
QY 562 AGCTAGTGGTCTGCTGCTTTATTCGACAGAGTTGGAAGAGAGAGAGAGAGAGAGAGAGAG 621
Db 349 AGCTAGTGGTCTGCTGCTTTATTCGACAGAGTTGGAAGAGAGAGAGAGAGAGAGAGAGAG 408
QY 622 CAGAGCTCCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 681
Db 409 CAGAGCTTACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 468
QY 682 TGCATCTCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 741
Db 469 TCTCTA-----GTGGCCCTTCACTGCGGAGGCTGTGCA 504
QY 742 AGTCAGAAATCTCCCGGACAGTGTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 801
Db 505 GCATGATACGAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 564
QY 802 AGCAGAGATTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 859
Db 565 TGGAACTCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 624
QY 860 ACTTATGCAAACTTACTATATCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 919
Db 625 ACNAGGCCCATGCTCTCTGGGCGCAGTGGTCTGTGTGACCT-----AAGG 672
QY 920 AAGTCAGAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 979
Db 673 ACATGGAGAGCTGGACATCAACCCCTGCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 732
QY 980 GTGAGCTCAGTGTGTCTCACTCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1038
Db 733 AAGCAGCCTTCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 792
QY 1039 ATTCATCTTCAGATGAGTATGATCTGCTTGGCCTGAGAGAGAGAGAGAGAGAGAGAGAG 1098
Db 793 TTATCTTCAGATGAGTATGATCTGCTTGGCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 852
QY 1099 CATATGAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1158
Db 853 CATATGAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 912
QY 1159 AACCATGGCATGAATTAATCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1218
Db 913 AACCATGGCATGAATTAATCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 972

```


QY	202	GCAGGCACTCGGTGCGCCATATGGGACCCGCGAGCGGCGCTTGGGAGCCCGAGCGTCC	261	1230	-----	1229	-----	1228	-----	1227	-----	1226	-----	1225	-----	1224	-----	1223	-----	1222	-----	1221	-----	1220	-----	1219	-----	1218	-----	1217	-----	1216	-----	1215	-----	1214	-----	1213	-----	1212	-----	1211	-----	1210	-----	1209	-----	1208	-----	1207	-----	1206	-----	1205	-----	1204	-----	1203	-----	1202	-----	1201	-----	1200	-----	1199	-----	1198	-----	1197	-----	1196	-----	1195	-----	1194	-----	1193	-----	1192	-----	1191	-----	1190	-----	1189	-----	1188	-----	1187	-----	1186	-----	1185	-----	1184	-----	1183	-----	1182	-----	1181	-----	1180	-----	1179	-----	1178	-----	1177	-----	1176	-----	1175	-----	1174	-----	1173	-----	1172	-----	1171	-----	1170	-----	1169	-----	1168	-----	1167	-----	1166	-----	1165	-----	1164	-----	1163	-----	1162	-----	1161	-----	1160	-----	1159	-----	1158	-----	1157	-----	1156	-----	1155	-----	1154	-----	1153	-----	1152	-----	1151	-----	1150	-----	1149	-----	1148	-----	1147	-----	1146	-----	1145	-----	1144	-----	1143	-----	1142	-----	1141	-----	1140	-----	1139	-----	1138	-----	1137	-----	1136	-----	1135	-----	1134	-----	1133	-----	1132	-----	1131	-----	1130	-----	1129	-----	1128	-----	1127	-----	1126	-----	1125	-----	1124	-----	1123	-----	1122	-----	1121	-----	1120	-----	1119	-----	1118	-----	1117	-----	1116	-----	1115	-----	1114	-----	1113	-----	1112	-----	1111	-----	1110	-----	1109	-----	1108	-----	1107	-----	1106	-----	1105	-----	1104	-----	1103	-----	1102	-----	1101	-----	1100	-----	1099	-----	1098	-----	1097	-----	1096	-----	1095	-----	1094	-----	1093	-----	1092	-----	1091	-----	1090	-----	1089	-----	1088	-----	1087	-----	1086	-----	1085	-----	1084	-----	1083	-----	1082	-----	1081	-----	1080	-----	1079	-----	1078	-----	1077	-----	1076	-----	1075	-----	1074	-----	1073	-----	1072	-----	1071	-----	1070	-----	1069	-----	1068	-----	1067	-----	1066	-----	1065	-----	1064	-----	1063	-----	1062	-----	1061	-----	1060	-----	1059	-----	1058	-----	1057	-----	1056	-----	1055	-----	1054	-----	1053	-----	1052	-----	1051	-----	1050	-----	1049	-----	1048	-----	1047	-----	1046	-----	1045	-----	1044	-----	1043	-----	1042	-----	1041	-----	1040	-----	1039	-----	1038	-----	1037	-----	1036	-----	1035	-----	1034	-----	1033	-----	1032	-----	1031	-----	1030	-----	1029	-----	1028	-----	1027	-----	1026	-----	1025	-----	1024	-----	1023	-----	1022	-----	1021	-----	1020	-----	1019	-----	1018	-----	1017	-----	1016	-----	1015	-----	1014	-----	1013	-----	1012	-----	1011	-----	1010	-----	1009	-----	1008	-----	1007	-----	1006	-----	1005	-----	1004	-----	1003	-----	1002	-----	1001	-----	1000	-----	999	-----	998	-----	997	-----	996	-----	995	-----	994	-----	993	-----	992	-----	991	-----	990	-----	989	-----	988	-----	987	-----	986	-----	985	-----	984	-----	983	-----	982	-----	981	-----	980	-----	979	-----	978	-----	977	-----	976	-----	975	-----	974	-----	973	-----	972	-----	971	-----	970	-----	969	-----	968	-----	967	-----	966	-----	965	-----	964	-----	963	-----	962	-----	961	-----	960	-----
----	-----	--	-----	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	------	-------	-----	-------	-----	-------	-----	-------	-----	-------	-----	-------	-----	-------	-----	-------	-----	-------	-----	-------	-----	-------	-----	-------	-----	-------	-----	-------	-----	-------	-----	-------	-----	-------	-----	-------	-----	-------	-----	-------	-----	-------	-----	-------	-----	-------	-----	-------	-----	-------	-----	-------	-----	-------	-----	-------	-----	-------	-----	-------	-----	-------	-----	-------	-----	-------	-----	-------	-----	-------	-----	-------	-----	-------	-----	-------	-----	-------	-----	-------	-----	-------

Db	2113	TCGAGGATACCGTAACTGTTGATAGATGGATATGTGAGAGTGTGGATGTGCAAT	2172	Location/Qualifiers
Qy	2293	GTCTTTGATCTGATGCTGTTGAAGACGTCTGTGGGTGTGTAACGGGAATAACTCAGCC	2352	1. .5610
Db	2173	GTCTTTGGATCTGATGCTGTGAAAGACGTCTGTGGGTGTGTAACGGGAATAACTCAGCC	2232	/organism="Homo sapiens"
Qy	2353	TGCACGATTCACAGGGGTCTCTACACAAGGACCAACACACCAACCAACGATTTATCATG	2412	/mol_type="genomic DNA"
Db	2233	TGCACGATTCACAGGGGTCTCTACACAAGGACCAACACACCAACCAACGATTTATCATG	2292	/db_xref="taxon:9606"
Qy	2413	GTACACATTCCTCTCTGAGCCCGGAGGTATCCGCACTCTATGAATGAAGCTCTCTACCTCC	2472	BASE COUNT 1233 a 1638 c 1537 g 1202 t
Db	2293	GTACACATTCCTCTCTGAGCCCGGAGGTATCCGCACTCTATGAATGAAGCTCTCTACCTCC	2352	ORIGIN
Qy	2473	TACATTTCTGTGCGCAATCCCTCAGAAAGTACTACCTGAATGGGCACTGGACCGTGGAC	2532	Query Match 64.9%; Score 1877.8; DB 6; Length 5610;
Db	2353	TACATTTCTGTGCGCAATCCCTCAGAAAGTACTACCTGAATGGGCACTGGACCGTGGAC	2412	Best Local Similarity 84.4%; Pred. No. 0;
Qy	2533	TGGCCCGGGCGGTACAAATTTTCGGGCACTACTTTCAGCTACAGACGGTCTTATATGAG	2592	Matches 289; Conservative 0; Mismatches 177; Indels 245; Gaps 6;
Db	2413	TGGCCCGGGCGGTACAAATTTTCGGGCACTACTTTCAGCTACAGACGGTCTTATATGAG	2472	
Qy	2593	CCCAGAACTTAATCGCTACTGACCAACCAACAGACACTGATGTGAGGCTGCTGTTT	2652	
Db	2473	CCCAGAACTTAATCGCTACTGACCAACCAACAGACACTGATGTGAGGCTGCTGTTT	2532	
Qy	2653	CAGGGAAGGAACCCGGTGTTCCTCGGAAATACTCCATGCTCGCTTGGGACCGAGAAG	2712	
Db	2533	CAGGGAAGGAACCCGGTGTTCCTCGGAAATACTCCATGCTCGCTTGGGACCGAGAAG	2592	
Qy	2713	CAGCCCTCCGCCAGCCAGCTACACITGGGCCATCGTGGCTCTGAGTGTCTCGTGTCC	2772	
Db	2593	CAGCCCTCCGCCAGCCAGCTACACITGGGCCATCGTGGCTCTGAGTGTCTCGTGTCC	2652	
Qy	2773	TGCGAGGGGG 2783		
Db	2653	TGCGAGGGGG 2663		
RESULT 5				
BD170080				
LOCUS				
DEFINITION				
Novel ADAMTS family polypeptide and gene encoding the same.				
ACCESSION				
BD170080				
VERSION				
BD170080.1 GI:27875892				
KEYWORDS				
WO 0231163-A/2.				
SOURCE				
Homo sapiens				
ORGANISM				
Homo sapiens				
REFERENCE				
AUTHORS				
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
TITLE				
Novel ADAMTS family polypeptide and gene encoding the same				
JOURNAL				
Patent: WO 0231163-A 2 18-APR-2002;				
KAZUSA DNA RESEARCH INSTITUTE, MITSUBISHI PHARMA CORP, OSAMU OHARA,				
TAKAHIRO NAGASE, NOBUO NOMURA, KAZUHIRO YANO, KOJI MURAKAMI, SHINICHIRO				
YASUDA, KOJI KANZAKI				
COMMENT				
OS Homo sapiens (human)				
PN WO 0231163-A/2				
PD 19-APR-2002				
PF 11-OCT-2001 WO 2001JP008913				
PR 11-OCT-2001 JP 01P 102905 FI				
OSAMU OHARA, TAKAHIRO NAGASE, NOBUO NOMURA, KAZUHIRO YANO, KOJI FI				
MURAKAMI,				
PI SHINICHIRO YASUDA, KOJI KANZAKI				
PC C12N15/57, C12N9/64, C12N5/00, C12N1/19, C12N1/21, C12N1/15, C07K16/				
PC 40, G01Q1/68,				
PC G01N33/50				
CC Novel ADAMTS family polypeptide and gene encoding the same FH				
Key				
Location/Qualifiers				
(770) . .(4444) .				
FT CDS				

[illegible]

Db	2702	GTTCGACTTCCTGCTGCTCTAGTGTGCGGAGCACAA	CAGCAGACGATTTCAGAGGGCGGCAC	2762
Qy	2113	TACAGGTGGAGCCCTTACACTCAAGTAGAAGATCAG	AGACTCAGGACTTATGCAGAACTCTACTCGTATC	2172
Db	2762	TACAGTGGAGCCCTTACACTCAGTAGAAGATCAG	AGACTCAGGACTTATGCAGAACTCTACTCGTATC	2821
Qy	2173	GCAGAGGATTTGATTTCTTTCTTTCTTTGTCAAA	TAAAGTCAAGATGSGAGCTCCATGCG	2232
Db	2822	GCAGAGGATTTGATTTCTTTCTTTGTCAAA	TAAAGTCAAGATGSGAGCTCCATGCG	2881
Qy	2233	TCGGAGGATAGCCGTAAATGTTGTATAGATGGGAT	TATGTGAGAGAGTGGGATGTGACAAAT	2292
Db	2882	TCGGAGGATAGCCGTAAATGTTGTATAGATGGGAT	TATGTGAGAGAGTGGGATGTGACAAAT	2941
Qy	2293	GTTCCTTGGATCTGATCTGTTTGAGAGAGCTCT	GTCGGGTGTGTAAACGGGAAATAACTCAGCC	2352
Db	2942	GTTCCTTGGATCTGATCTGTTTGAGAGAGCTCT	GTCGGGTGTGTAAACGGGAAATAACTCAGCC	3001
Qy	2353	TGCACGATTACAGGGGTCTCTACACAAGACACCA	CACACCAACCAACCAAGTATTATCATGTG	2412
Db	3002	TGCACGATTACAGGGGTCTCTACACAAGACACCA	CACACCAACCAACCAAGTATTATCATGTG	3061
Qy	2413	GTACACATTCCTTTCTGGAGCCGAGGTATCCGG	ACATCTATGAAATGAAAGTCTCTACCTCC	2472
Db	3062	GTACACATTCCTTTCTGGAGCCGAGGTATCCGG	ACATCTATGAAATGAAAGTCTCTACCTCC	3121
Qy	2473	TACATTTCTGTGCGCAATCCCTCTCAGAGGTA	CTATCTGATATGGGCACTGACACGTGTGAC	2532
Db	3122	TACATTTCTGTGCGCAATCCCTCTCAGAGGTA	CTATCTGATATGGGCACTGACACGTGTGAC	3181
Qy	2533	TGGCCCGGCGGTCACAAATTTCTGGAGCACTCT	TCGACTCAGACGGTCCGATTAATATGAG	2592
Db	3182	TGGCCCGGCGGTCACAAATTTCTGGAGCACTCT	TCGACTCAGACGGTCCGATTAATATGAG	3241
Qy	2593	CCCGGAGACTTAATCGCTACTGGACACCAACCA	ACGACACACTGATTTGTGGAGCGTCTGTTT	2652
Db	3242	CCCGGAGACTTAATCGCTACTGGACACCAACCA	ACGACACACTGATTTGTGGAGCGTCTGTTT	3301
Qy	2653	CAGGGAGGACACCCGGGTGTGCTCGGAAATACT	TCCATGCCTCGCTTGGGACACGAGAG	2712
Db	3302	CAGGGAGGACACCCGGGTGTGCTCGGAAATACT	TCCATGCCTCGCTTGGGACACGAGAG	3361
Qy	2713	CAGCCCTCGCCACAGCCAGCTACACTTGGGCG	CAATCGTCCGCTGAGTGTCTCGGTGCC	2772
Db	3362	CAGCCCTCGCCACAGCCAGCTACACTTGGGCG	CAATCGTCCGCTGAGTGTCTCGGTGCC	3421
Qy	2773	TGCGGAGGGGG	2783	
Db	3422	TGCGGAGGGGG	3432	

RESULTS

RESULT 6
BD1700A3

LOCUS
S1301

DEFINITION

DEFINITION
ACCESSION

REVISION
VERSION

KEYWORDS

SOURCE

ORGANI

REFERENCES

AUTHOR:

1
1
1
1
1

TITLE

JOURNALS

COMMENT

COMMENT

BD170083 8435 bp DNA linear PAT 17-JAN-2003
Novel ADAMTS family polypeptide and gene encoding the same.

NOVEL ADAMIS FAMILY POLYPERPANE AND GELRE ENCOATING THE SAME.
BD170083

BD170083.1 GI:27875895

WO 0231163-A/5.

synthetic construct

synthetic construct

artificial sequences.

1 (bases 1 to 8435)

Ohara, O., Nagase, T., Nomura, N., Yano, K., Murakami, K., Yasuda, S. and

Kanzaki, K.

Novel ADAMTS family polypeptide and gene encoding the same

Patent: WO 0231163-A 5 18-APR-2002;

KAZUSA DNA RESEARCH INSTITUTE, MITSUBISHI PHARMA CORP., OSAMU OHARA,

TAKAHIRO NAGASE, NOBUO NOMURA, KAZUHIRO YANO, KOJI MURAKAMI, SHINICHIRO YASUDA, KOJI KANUNAY

YASUDA, KOJI KANZAKI
OS Artificial Cement

US Artificial Sequence
 PW 0231157-2/5

PN WO 0231163-A/5
PD 18-APR-2003

```

PR 11-OCT-2000 JP OOP 311309,02-APR-2001 JP OIP 102905 PI
OSAMU OHARA,TAKAHIRO NAGASE,NOBUO NOMURA,KAZUHIRO YANO,KOJI PI
MURAKAMI,
PI SHINICHIRO YASUDA,KOJI KANZAKI
PC C12N15/57,C12N9/64,C12N5/00,C12N1/19,C12N1/21,C12N1/15,C07K16/
PC 40,C12Q1/68,
PC GOIN33/50
CC Description of Artificial Sequence:base sequence of the CC
plasmid
CC pFastBacl-HT-PJ01256
FH Key Location/Qualifiers
FT source 1..8435
FT /organism='Artificial sequence'
FEATURES
source
Location/Qualifiers
1..8435
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
BASE COUNT 2072 a 2174 c 2149 g 2040 t
ORIGIN
Query Match 64.9%; Score 1877.8; DB 6; Length 8435;
Best Local Similarity 84.4%; Pred. No. 0;
Matches 2289; Conservative 0; Mismatches 177; Indels 245; Gaps 6;
QY 202 GCAGGCACTCGCTGGCCATCGGACCCGACGCGCGAGCGCTGGAGCCCGAGCGTCC 261
DB 4120 GCAGGCACTCGCTGGCCATCGGACCCGACGCGCGAGCGCTGGAGCCCGAGCGTCC 4179
QY 262 GCCTCTCTCCACCCGCGGACCCGCGGCGCTGGATGGAAAAGCGCGGACATGGATGAA 321
DB 4180 GCCTCTCTCCACCCGCGGACCCGCGGCGCTGGATGGAAAAGCGCGG----- 4226
QY 322 CTTGGAAACCATCGTCTCAGCAACTAACAACAGGAAACAGAAACCAACATCGCATGT 381
DB 4227 ----- 4226
QY 382 CTCACATCATATGACCTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 441
DB 4227 -----ATATGACCTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4279
QY 442 CATGAATATCATCAATCATGCGCGGAGAGAGCATGGCTGGCTGGAGTTGAGTT 501
DB 4280 CATGAATATCATCAATCATGCGCGGAGAGAGCATGGCTGGAGTTGAGTTGAGTT 4339
QY 502 CTTCACTTGGCTGAAAGGCGCCGAGGAGACTTCCATGAGTCTGAGGACTTCCAGC 561
DB 4340 CTTCACTTGGCTGAAAGGCGCCGAGGAGACTTCCATGAGTCTGAGGACTTCCAGC 4399
QY 562 AGCCTAGTGGCTCTGGCTTTATGTGAGAGGTTGGGAAAGAGAGGACTAGTCTGTG 621
DB 4400 AGCCTAGTGGCTCTGGCTTTATGTGAGAGGTTGGGAAAGAGAGGACTAGTCTGTG 4459
QY 622 CAGACTTTACCGCGAGAGGCTCTGTGTTCTATCAGGCTCTTCCGATCAGCAGAAC 681
DB 4460 CAGACTTTACCGCGAGAGGCTCTGTGTTCTATCAGGCTCTTCCGATCAGCAGAAC 4519
QY 682 TGGCATCGCATGGAGGAGAGTTCTGTGAGGCGTCCACTCGACTCTGAGCTCTGCAC 741
DB 4520 TCCCTA-----GTGGCCCTTTCACCTCTCCAGGCTGTCA 4555
QY 742 AGTCAGAAATGTCCCGGAGCAGTGTGACTTCCGCTGCTGCTGCTGCTGCTGCTGCTG 801
DB 4556 GGCATGATGAGAACAGAGAGGAGCAGTACTTCTTAGGCGACTTCTTCACTCTCA 4615
QY 802 AGCAGAGATTCAGAGGCGGCGCTACAGTGG--AGCCTTACACTCAGTATGAGCGG 859
DB 4616 TGGAAACTTCGCGAGAGCTGCCAGGCGCTGCCATCCCACTGCTACAGAGATCC 4675
QY 860 ACTTATGAAACTCTACTGTATGTCGAGAGAGATTTGATTTCTTTCTTTCTCAATA 919
DB 4676 ACAGAGCCCATCTCTCTGGGCGCAGTGAAGCTCTGTGGTCTCT-----AAGG 4723

```

```

QY 920 AAGTCAAGAGTGGGATCTCATCTCGAGGATAGCCGTAATGTTTGTATATGGGAT 979
DB 4724 AAGTCAAGAGTGGGATCTCATCTCGAGGATAGCCGTAATGTTTGTATATGGGAT 4783
QY 980 GTGAGCTCAGTGTGGTCTCCAGATGTC- GCAGTGGCCCGAGGCTCCCAAGAGACCTC 1038
DB 4784 AAGCAGCATTTCTGTGGAAGCCGAGAAATACATGCCCGAGCTCCCAAGAGACCTC 4843
QY 1039 TTTCTTGTGCGAGTACGATATGCTTGTACGGCATAGCGCTCTCTTCTGAGGTCC 1098
DB 4844 TTTCTTGTGCGAGTACGATATGCTTGTACGGCATAGCGCTCTCTTCTGAGGTCC 4903
QY 1099 CATGAAATGAGAACTGAGCTGGAGACTTGGTGGTGGCGACAAAAGATGATGCAA 1158
DB 4904 CATGAAATGAGAACTGAGCTGGAGACTTGGTGGTGGCGACAAAAGATGATGCAA 4963
QY 1159 AAGCATGGCCCTGAAAATATCATCACTTGTGCTGCTGCTGCTGCTGCTGCTGCT 1218
DB 4964 AAGCATGGCCCTGAAAATATCATCACTTGTGCTGCTGCTGCTGCTGCTGCTGCT 5023
QY 1219 TTTATCAAGG----- 1229
DB 5024 TTTATCAAGGATGGAACATATGAGGAGAAACATCAACATGCAATTTGAGTCTGATTCTT 5083
QY 1230 ----- 1229
DB 5084 CTGAGAGATGAAACAGCCAGGACTGGTGATAGTCAACACGACGACACACCTTAAGTAGC 5143
QY 1230 -----TGGATTTGATGGGAAAGATGGGATCGTCAATGACACGACGATC 1272
DB 5144 TTTCTGCGAGTGGCAGTCTGGATTTGATGGGAAAGATGGGACTCGTCAATGACACGATC 5203
QY 1273 TTTCTGAGTCTGCTGCTGATATGCTTCTGGAAGATGAGCTCTGTCGACACTTTGGGATTT 1332
DB 5204 TTTCTGAGTCTGCTGCTGATATGCTTCTGGAAGATGAGCTCTGTCGACACTTTGGGATTT 5263
QY 1332 GCACCCATGAGTGGATTTGATTAATATCGCAGCTGACAGATTAATGAAGATACAGGT 1392
DB 5264 GCACCCATGAGTGGATTTGATTAATATCGCAGCTGACAGATTAATGAAGATACAGGT 5323
QY 1393 CTTGGACTGGCTTACCAATTTGCCATGATCTGGAACAATTTGGCATGATTCATGAT 1452
DB 5324 CTTGGACTGGCTTACCAATTTGCCATGATCTGGAACAATTTGGCATGATTCATGAT 5383
QY 1453 GGAGAGGAGACATGTTGTAABAAAGTCGAGGAGCAACATCATGTCCCTACATTTGGCAGGA 1512
DB 5384 GGAGAGGAGACATGTTGTAABAAAGTCGAGGAGCAACATCATGTCCCTACATTTGGCAGGA 5443
QY 1513 GCGAATGGAGTCTTCTCTGCTGATCCCTGAGCCGCGAGTATCTACAGAAATTTCTAAGC 1572
DB 5444 GCGAATGGAGTCTTCTCTGCTGATCCCTGAGCCGCGAGTATCTACAGAAATTTCTAAGC 5503
QY 1573 ACCTCTCAAGTATCTGCTTCTGCTGATCAGCCGAGGCTTGAAGGATTAACAGTATCTC 1632
DB 5504 ACCTCTCAAGTATCTGCTTCTGCTGATCAGCCGAGGCTTGAAGGATTAACAGTATCTC 5563
QY 1633 GAGAAATTTCCGAGGAGATTAATGATGCAACACACAGTCAAGTGGAGTTCGGAGAG 1692
DB 5564 GAGAAATTTCCGAGGAGATTAATGATGCAACACACAGTCAAGTGGAGTTCGGAGAG 5623
QY 1693 AAGGCAAGCTCTCAGTCTGCTTTTAAAGGAGCATCTGTAAGCTCTGTTGGTCCAT 1752
DB 5624 AAGGCAAGCTCTCAGTCTGCTTTTAAAGGAGCATCTGTAAGCTCTGTTGGTCCAT 5683
QY 1753 GGTATTGGAGGAGTCTGAGCATTAATTTATGCGAGGAGAGAGGAGCATTTTGGG 1812
DB 5684 GGTATTGGAGGAGTCTGAGCATTAATTTATGCGAGGAGAGAGGAGCATTTTGGG 5743
QY 1813 CATGACATGCTGCTGGGAGGAGCAGTGTGAAATATGTTGATGAGGCCCCCAAGCCC 1872
DB 5744 CATGACATGCTGCTGGGAGGAGCAGTGTGAAATATGTTGATGAGGCCCCCAAGCCC 5803
QY 1873 ACCCATGGCCACTGGTGGGACTGGTCTTCTTGGTCCCATGCTCCAGGACCTGCGGAGG 1932

```


2773 TGC GGAGGGGG 2783

Db	6774	TGCGAGGGGG	6784
RESULT 8			
LOCUS	AR229456	3675 bp	DNA linear PAT 20-DEC-2002
DEFINITION	Sequence 3 from patent US 6449388.		
ACCESSION	AR229456		
VERSION	AR229456.1	GI:27268950	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 3675)		
AUTHORS	Fridde,C.J. and Hilbun,E.		
TITLE	Human processes and polynucleotides encoding the same		
JOURNAL	Patent: US 6449388-A 3 10-SEP-2002;		
FEATURES	Location/Qualifiers		
source	1..3675		
BASE COUNT	873 a 1008 c 1020 g 774 t		
ORIGIN	/organism="unknown"		
Query Match	64.8%; Score 1874.6; DB 6; Length 3675;		
Best Local Similarity	84.4%; Fred. No. 0;		
Matches 2287; Conservative	0; Mismatches 179; Indels 245; Gaps 6;		
QY	202	GCAGCCACTCGTGGGCCCATGGAACCCGACGCGCAGCGCTGTGGAGCCGAGCGTGCCC	261
DB	69	GCRGGCACCTGCTGGCGCCATGGAACCCGACGCGCAGCGCTGTGGAGCCGAGCGTGCCC	128
QY	262	CGCTCTCTCCACCCTGGAGCGCGCGGCTGTGATGGAAAGGCGCGGACATGAGTAAGA	321
DB	129	CGCTCTCTCTCCACCCTGGAGCGCGCGGCTGTGATGGAAAGGCGCGG	175
QY	322	CTCGAACAACATCGTTCTTCAGCAAACTAACACAGGAACAGAAAACCAACACTGCATGTT	381
DB	176	-----	175
QY	382	CTCACTCAATATGACTGTGTCTTCGCTACAGAGTTTACACAGGGGGGATTAGTGTTCC	441
DB	176	-----AATATGACTGTGTCTTCGCTACAGAGTTTACACAGGGGGGATTAGTGTTCC	228
QY	442	CATGAATCATGCACCATCAGCGCGGAGAGAGTAGTGCGGCTGTGCGAGTTGAGTCT	501
DB	229	CATGAATCATGCACCATCAGCGCGGAGAGAGTAGTGCGGCTGTGCGAGTTGAGTCT	288
QY	502	CTTCACCTTCGGCTGAAAGGCCCCAGCGCGACATTCACATGGAATCTGAGGACTCCACG	561
DB	289	CTTCACCTTCGGCTGAAAGGCCCTCAGGCGACATTCACATGGAATCTGAGGACTCCACG	348
QY	562	AGCTAGTGGCTCTGGCTTTATTTGTGCGAGCGTTGGAAAGACAGGCGACATAAGTCTGTG	621
DB	349	AGCTAGTGGCTCTGGCTTTATTTGTGCGAGCGTTGGAAAGACAGGCGACATAAGTCTGTG	408
QY	622	CNAGCTTTACCGCAGGAGGACTCTGTGTTCTATCAAGGCTCTTTGCGATCACAGAAAC	681
DB	409	CNAGCTTTACCGCAGGAGGACTCTGTGTTCTATCAAGGCTCTTTGCGATCACAGAAAC	468
QY	682	TGCGCATTCGATGAGGGAAGTTCTGTGAGGGCTCCACCTCGCACTCTGAAGCTCTCGAAC	741
DB	469	TCCTCA-----GTGGCCCTTTCAACCTCGCCAAGGCTTGCA	504
QY	742	AGTCAGAAATGTCCCGGAGCAGTGTGATCTTCGCTGTGCTCAGTGTGCGGAGCACAC	801
DB	505	GGCATGATACGAACAGMAGGCGAGATTACTTCTTAAGGCCACTTCTTTACACCTCTCA	564
QY	802	AGCAGACGATTACAGAGGGGGCAGCTACAGATGG--AAGCCTTACACTCAAGTAGAGCGG	859
DB	565	TGGAATCTGGCGAGAGCTGCCCAAGGCGAGCTGCGCATCCACGCTACTGTACAGAGATCC	624
QY	860	ACTTATGCAAACTCTACTGTATTCGCAAGAGGATTGATTTCTTTTTCTTTTGTGTCAAATA	919

Db 505 GCGATGATACGACACAGAGGCGAGATTACTTCTTAAGGCCAATTCCTTCAACCTCTCA 564
Qy 802 AGCTAGAGATTCCAGGGCGGCACTACAGTGG--NAGCTTACACTCAAGTAGAGACGG 859
Db 565 TGGAAATCTCGGAGAGCTGCCAGGAGCTCGCATCCACAGTACTGTACAAGAGATCC 624
Qy 860 ACTTATGCAAACTACTATATCGCAGAGGATTTGATTTCTTTCTTTCTTGCAATA 919
Db 625 ACAGCCCAAGCTCTTGGGCGAGTGAGTCTGGTGACCTC-----AAGG 672
Qy 920 AAGTCAAGATGGGATCTCATCTCGGAGGATGAGCTTAATGTTGTATAGATGGGATAT 979
Db 673 AATGGGAGCTGACATCAACGCCCTGCACAGCAGGACCTTCGCTGGGACTGCCACAA 732
Qy 980 GTGAGCTCAGTGTGTCTCATCTCGGAGGATGAGCTTAATGTTGTATAGATGGGATAT 1038
Db 733 AAGCAGATTTCTGTGGAAGACGCAAGAAATACATGCTCCGCTGGGACTGCCACAA 792
Qy 1039 TTCTCTCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1098
Db 793 TTCTCTCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 852
Qy 1099 CATAGAAATGAGAACTGAACTGAGGAGAGCTTGGTGTGTGTGTGTGTGTGTGTGTGTGT 1158
Db 853 CATAGAAATGAGAACTGAACTGAGGAGAGCTTGGTGTGTGTGTGTGTGTGTGTGTGTGT 912
Qy 1159 AACCATGGCAGTAAATATCAACCTACGCTGCTACGATATCTCAACATGATGATGATGAT 1218
Db 913 AACCATGGCAGTAAATATCAACCTACGCTGCTACGATATCTCAACATGATGATGATGAT 1272
Qy 1219 TTATTTCAAGA----- 1229
Db 973 TTATTTCAAGATGGAACATAGGAGGAAACATCAACATTCGAATGTAGTGTGATTTCTT 1032
Qy 1230 ----- 1229
Db 1033 CTAGAGATGACAGCAGGAGCTGTGTAAAGTACCAAGCAGACACACCTTAAAGTAGC 1092
Qy 1230 -----TGGATTTGATGGGAAGATGGGACTCGTCATGACCAAGCAGC 1272
Db 1093 TTCTCGAGTGGGAGCTCTGGATTTGATGGGAAAGATGGGACTCGTCATGACCAAGCAGC 1152
Qy 1273 TTACTGACTGCTGATATATGTTCTTGGAGATGAGCCTGTGACACTTTGGGATTT 1332
Db 1153 TTACTGACTGCTGATATATGTTCTTGGAGATGAGCCTGTGACACTTTGGGATTT 1212
Qy 1333 GCACCATAGTGGATGTGATTAATATGCGAGCTGACAGATTAATAGAGATACAGT 1392
Db 1213 GCACCATAGTGGATGTGATTAATATGCGAGCTGACAGATTAATAGAGATACAGT 1272
Qy 1393 CTTCGACTGGCCTTACCATGCTGAGCTTGGACACAACTTTGGCATGATTCATGAT 1452
Db 1273 CTTCGACTGGCCTTACCATGCTGAGCTTGGACACAACTTTGGCATGATTCATGAT 1332
Qy 1453 GGGAGAGGACATGTGTAAGAGTCCGAGGCAACATCATGTCCCTACATGGCAGGA 1512
Db 1333 GGGAGAGGACATGTGTAAGAGTCCGAGGCAACATCATGTCCCTACATGGCAGGA 1392
Qy 1513 GCGATAGGATCTTCTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1572
Db 1453 GCGATAGGATCTTCTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1512
Qy 1633 GAGAAATTCGCGAGGATTAATATGATGATGATGATGATGATGATGATGATGATGATGAT 1692
Db 1513 GAGAAATTCGCGAGGATTAATATGATGATGATGATGATGATGATGATGATGATGATGAT 1572
Qy 1693 AAGAGCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1752
Db 1573 AAGAGCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1632

1753 GGTATTGGAAGAAATGTGAGACTAAATTTATGCCAGCAGCAGAGGCAAAATTTGTGGG 1812
1633 GGTATTGGAAGAAATGTGAGACTAAATTTATGCCAGCAGCAGAGGCAAAATTTGTGGG 1692
1813 CATGACATGTTGGTSCCGGGGAGGACAGTGTGTGAAATATGTTGATGAAAGGCCCAAGCCC 1872
1693 CATGACATGTTGGTSCCGGGGAGGACAGTGTGTGAAATATGTTGATGAAAGGCCCAAGCCC 1752
1873 ACCCATGGCCACTGTTGGTGGACTGTTCTTTGGTCCCATGCTCCAGACCTTGGGGAGG 1932
1753 ACCCATGGCCACTGTTGGTGGACTGTTCTTTGGTCCCATGCTCCAGACCTTGGGGAGG 1812
1933 GGAGTATCTCATAGAGTGGCTCTGCAACAAACCCCAAGCCATCGCATGGAGGGAAGTTTC 1992
1813 GGAGTATCTCATAGAGTGGCTCTGCAACAAACCCCAAGCCATCGCATGGAGGGAAGTTTC 1872
1993 TGTGAGGGCTCCACTGSCACTCTGAGCTCTGCAACAGCTCAGAAATGTTCCCGGGAGCAGT 2052
1873 TGTGAGGGCTCCACTGSCACTCTGAGCTCTGCAACAGCTCAGAAATGTTCCCGGGAGCAGT 1932
2053 GTTCACTTCGTTGCTGCTCAGTGTGCGAGCACAACAGCAGACGATTGAGAGGGCGGCAC 2112
1933 GTTCACTTCGTTGCTGCTCAGTGTGCGAGCACAACAGCAGACGATTGAGAGGGCGGCAC 1992
2113 TACAAGTGAAGCTTTACACTCAAGTAGAAGATCAGGACTTATGCAAACTCTTACTGTATC 2172
1993 TACAAGTGAAGCTTTACACTCAAGTAGAAGATCAGGACTTATGCAAACTCTTACTGTATC 2052
2173 GCAGAGGATTTGATTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 2232
2053 GCAGAGGATTTGATTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 2112
2233 TCGAGAGTAGCCGTAAATGTTGTATAGATGGGATATGTGAGAGAGTTGGATGTGACAAAT 2292
2113 TCGAGAGTAGCCGTAAATGTTGTATAGATGGGATATGTGAGAGAGTTGGATGTGACAAAT 2172
2293 GTCTCTTGATCTGTGCTGTGAGAGAGTCTGTGGGGTGTGTAAACGGGAAATAACTCAGCC 2352
2173 GTCTCTTGATCTGTGCTGTGAGAGAGTCTGTGGGGTGTGTAAACGGGAAATAACTCAGCC 2232
2353 TGCAAGATTCAAGGGGCTCTTACACCAAGCACCACCAACCAACCAAGTATTATCACAATG 2412
2233 TGCAAGATTCAAGGGGCTCTTACACCAAGCACCACCAACCAACCAAGTATTATCACAATG 2292
2413 GTCAACATTCCTTTCTGAGCCCGGAGTATCCGCTATGAAATGAAAGCTCTTACTCTCC 2472
2293 GTCAACATTCCTTTCTGAGCCCGGAGTATCCGCTATGAAATGAAAGCTCTTACTCTCC 2352
2473 TACATTTCTGTGCGCAATGCTCTCAGAGGATCTACTTGAATGGGCACTGGAGCGTGAC 2532
2353 TACATTTCTGTGCGCAATGCTCTCAGAGGATCTACTTGAATGGGCACTGGAGCGTGAC 2412
2533 TGCGCCGGCGGTAACAAATTTTGGGCACTACTTTGCACTACAGCAGGCTCTATTAATGAG 2592
2413 TGCGCCGGCGGTAACAAATTTTGGGCACTACTTTGCACTACAGCAGGCTCTATTAATGAG 2472
2593 CCGGAGAACTTAATCGCTACTGAGCAACCAACGAGACACTGATTTGTGGAGCTGTGTTT 2652
2473 CCGGAGAACTTAATCGCTACTGAGCAACCAACGAGACACTGATTTGTGGAGCTGTGTTT 2532
2653 CAGGAGAGAAACCCGGGTGTTCCTGGGAATATCTCCATGCTCGCTTGGGGAACGAGAG 2712
2533 CAGGAGAGAAACCCGGGTGTTCCTGGGAATATCTCCATGCTCGCTTGGGGAACGAGAG 2592
2713 CAGGCCCCCTGCGCAGCCAGCTACACTTTGGGCCATCGTGGCTCTGAGTGTCTCGTGTGCC 2772
2593 CAGGCCCCCTGCGCAGCCAGCTACACTTTGGGCCATCGTGGGCCATCGTGGCTCTGAGTGTCTCGTGTGCC 2652
2773 TCGGAGGGGG 2783
2653 TCGGAGGGGG 2663

RESULT 10
AR229457 LOCUS 4042 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 5 from patent US 6448388.
AR229457 ACCESSION
AR229457.1 GI:27268951
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4042)
Fridlie-C.J. and Hilbun, E.
AUTHORS Human proteases and polynucleotides encoding the same
TITLE Patent: US 6448388-A 5 10-SEP-2002;
JOURNAL Location/Qualifiers
FEATURES
source 1..4042
/organism="unknown"
BASE COUNT 947 a 1117 c 1130 g 848 t
ORIGIN
Query Match 64.8%; Score 1874.6; DB 6; Length 4042;
Best Local Similarity 84.4%; Pred. No. 0;
Matches 2287; Conservative 0; Mismatches 179; Indels 245; Gaps 6;
QY 202 GCAGCACCCTGCGCGCCATGCGACCCGCGAGCGGCGCTGCGAGCCCGGAGGTCCTCC 261
Db 167 GCAGCACCCTGCGCGCCATGCGACCCGCGAGCGGCGCTGCGAGCCCGGAGGTCCTCC 226
QY 262 GCGTCCTCTCCACCCGCGAGCGGCGCGCGCTGCGAGGCGGCGGCGAGTGA 321
Db 227 GCGTCCTCTCCACCCGCGAGCGGCGCGCGCTGCGAGGCGGCGGCGGCGG 273
QY 322 GCTGGAACCATCTGCTCTCTGAGCAACTACACAGGACGAAACCAACACCTGCTGTT 381
Db 274 ----- 273
QY 382 CTCACCTCATATGACCTGCTCTCCCTGACGAGTGTGACACAGCGGCGGATCTGCTGCC 441
Db 274 -----AATGAGCTGGTCTCTGCTGACGAGTGTGACACAGCGGCGGATCTGCTGCC 326
QY 442 CATGAATATGACATGACGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGTCT 501
Db 327 CATGAATATGACATGACGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGTCT 386
QY 502 CTCACCTCTGCTGAAAGGCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 561
Db 387 CTCACCTCTGCTGAAAGGCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 446
QY 562 AGCCTAGTGGCTCTGCGCTTATGTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 621
Db 447 AGCCTAGTGGCTCTGCGCTTATGTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 506
QY 622 CAGACTTATGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 681
Db 507 CAGACTTATGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 566
QY 682 TCGCCTGCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 741
Db 567 TCTCTA-----GTGGCGCTTTCACCTGCGCAAGGCTGTGCA 602
QY 742 AGTCAGAAATGTCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 801
Db 603 GGCATGATACGACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 662
QY 802 AGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 859
Db 663 TGGAACTCCGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 722
QY 860 ACTTATGCAACTCTACTGATATCGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 919
Db 723 ACAGAGCCCACTCTCTCTGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 770

QY 920 AAGTCAAGATGGGAGCTCCATGCTCGGAGGATAGCCGTAATGTTGTTATAGATGGGATAT 979
Db 771 ACATGGGAGTGGGACATCAACCCCTGACAGAGGACGCTTCGCTCGGACTGCCACAA 830
QY 980 GTAGGCTCAGTGTGTGCTCCATCTGCG-GGACATGCGCCAGCCTCCGAGGAGAGCTC 1038
Db 831 AAGCAGCATTTCTGTGGAAGAGCGGAGAAATACATGACCCGAGCTCCGAGGAGAGCTC 890
QY 1039 TTCACTCTCCAGATGAGATATAGTCTTCCTACGGATAGAGCTCTCTCTGAGGTC 1098
Db 891 TTCACTCTCCAGATGAGTATAGTCTTCCTACGGATAGAGCTCTCTCTGAGGTC 950
QY 1099 CATGAATATGAAGAACTGAAAGCTGAGAGCTTGTGTGTGTCGACAAAGATGATGCA 1158
Db 951 CATGAATATGAAGAACTGAAAGCTTGTGTGTGTCGACAAAGATGATGCA 1010
QY 1159 AACCATGGCCATGAAATATACCACTACGTCCTACGATCTCAACATGCTATGCT 1218
Db 1011 AACCATGGCCATGAAATATACCACTACGTCCTACGATCTCAACATGCTATGCT 1070
QY 1219 TTATTCMAAGA----- 1229
Db 1071 TTATTCMAAGATGGAAACATATAGGAGGAAACATCAACATTTGATGAGTCTGATTC 1130
QY 1230 ----- 1229
Db 1131 CTAGAGATGAAACAGCCGAGCTGTGTATAGTCAACAACGAGACACACCTTAAGTAGC 1190
QY 1230 -----TGCAATGAGGAGGAGAGTGGAGCTCTGATGACACAGCCGATC 1272
Db 1191 TTCTGCGAGTGGCAGCTCTGATTTGTGGGAAAGATGGGATCTCTGATGACACGCTC 1250
QY 1273 TTAATGCTCTGCTGAGTATGATCTCTGGAAGATGAGGCTCTGACATCTGCGATTT 1332
Db 1251 TTAATGCTCTGCTGAGTATGATCTCTGGAAGATGAGGCTCTGACATCTGCGATTT 1310
QY 1333 GCACCAATGAGTGAATCTGATATATATGCGAGTGCACCATTAATGATGATGATGAT 1392
Db 1311 GCACCAATGAGTGAATCTGATATATATGCGAGTGCACCATTAATGATGATGATGAT 1370
QY 1393 CTTCGAGCTGCTTACATATGCGCAGTCTGCGAGACATCTTGGCATGATTCATGAT 1452
Db 1371 CTTCGAGCTGCTTACATATGCGCAGTCTGCGAGACATCTTGGCATGATTCATGAT 1430
QY 1453 GAGAGAGGAGATCTCTTAAAGTCCGAGGCGAGGACATCTGCTCCCTCATGATGCGAGA 1512
Db 1431 GAGAGAGGAGATCTCTTAAAGTCCGAGGCGAGGACATCTGCTCCCTCATGATGCGAGA 1490
QY 1513 GCGAATGAGTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1572
Db 1491 GCGAATGAGTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1550
QY 1573 ACCCTGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1632
Db 1551 ACCCTGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1610
QY 1633 GAGAAATTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1692
Db 1611 GAGAAATTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1670
QY 1693 AAGGCGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1752
Db 1671 AAGGCGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1730
QY 1753 GATATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1812
Db 1731 GATATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1790
QY 1813 CATGAGCTGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1872
Db 1791 CATGAGCTGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1850
QY 1873 ACCCATGGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1932

Db 1951 ACCATGGCCACTGGTGGAGTGGTCTTCTTGGTCCCGATGCTCCAGGACCTCGGAGGG 1910

Qy 1933 GGNATGATCTATAGAGTGGCCCTCTGACCAACCCCAAGGCATCCGATGAGGGAAGTTC 1992

Db 1911 GGNATGATCTATAGAGTGGCCCTCTGACCAACCCCAAGGCATCCGATGAGGGAAGTTC 1970

Qy 1993 TGTGAGGGCTCCACTGCGACTCTGAAGCTCTGCAACAGTCAAGAAATGTCCTCCGGACAGT 2052

Db 1971 TGTGAGGGCTCCACTGCGACTCTGAAGCTCTGCAACAGTCAAGAAATGTCCTCCGGACAGT 2030

Qy 2053 GTTGACTTTCGGTGGTCTGCTGAGTGGCGGACCAACAGCAGATTCAGAGGGGGCAC 2112

Db 2031 GTTGACTTTCGGTGGTCTGCTGAGTGGCGGACCAACAGCAGATTCAGAGGGGGCAC 2090

Qy 2113 TACAGTGGAGGCTTACCTCACTCAAGTACAGAGATCAGGACTTATGCAAACTCTACTGTATC 2172

Db 2091 TACAGTGGAGGCTTACCTCACTCAAGTACAGAGATCAGGACTTATGCAAACTCTACTGTATC 2150

Qy 2173 GCAGAGGATTTGATTTCTCTCTTTCTTCTTCAATTAAGTCAAGATGGGACTCCATGC 2232

Db 2151 GCAGAGGATTTGATTTCTCTCTTTCTTCTTCAATTAAGTCAAGATGGGACTCCATGC 2210

Qy 2233 TGGGAGATAGCCGTAATGTTCTATAGATGGGATATGAGAGAGTTGGATGACAAAT 2292

Db 2211 TGGGAGATAGCCGTAATGTTCTATAGATGGGATATGAGAGAGTTGGATGACAAAT 2270

Qy 2293 GTCTTGTGATCTGATCTGTTGGAAGACCTCTGTGGGGTGTAAAGGGAATTAATCAGCC 2352

Db 2271 GTCTTGTGATCTGATCTGTTGGAAGACCTCTGTGGGGTGTAAAGGGAATTAATCAGCC 2330

Qy 2353 TGCAGGATTCACAGGGGTCTCTACCAAGCACCAACACCAACCAACCAACCAATATACATG 2412

Db 2331 TGCAGGATTCACAGGGGTCTCTACCAAGCACCAACACCAACCAACCAACCAATATACATG 2390

Qy 2413 GTCACATCTCTCTGAGGCGCGGAGTATCGCATCTATGAATGAATGAATGCTCTCACTCC 2472

Db 2391 GTCACATCTCTCTGAGGCGCGGAGTATCGCATCTATGAATGAATGAATGCTCTCACTCC 2450

Qy 2473 TACATTTCTGTGGCAATAGCCCTCTAGAAAGTACTACCTGAATGGGCACTGGACCGTGAC 2532

Db 2451 TACATTTCTGTGGCAATAGCCCTCTAGAAAGTACTACCTGAATGGGCACTGGACCGTGAC 2510

Qy 2533 TGGCCCGGCGGTACAAATTTTCGGGCACACTTTTCGACTACAGACGGTCTCTAATGAG 2592

Db 2511 TGGCCCGGCGGTACAAATTTTCGGGCACACTTTTCGACTACAGACGGTCTCTAATGAG 2570

Qy 2593 CCGGAACTTAATCGCTACTGGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 2652

Db 2571 CCGGAACTTAATCGCTACTGGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 2630

Qy 2653 CAGGAGAGGAACCCGGGTGTGGGATPACTCCATGCTCTGCTTGGGACCGAGAG 2712

Db 2631 CAGGAGAGGAACCCGGGTGTGGGATPACTCCATGCTCTGCTTGGGACCGAGAG 2690

Qy 2713 CAGCCCGCTGCGCAGCCAGTCTACCTTGGGCACTGGTGGCTCTGAGTGTCTCGGTGCC 2772

Db 2691 CAGCCCGCTGCGCAGCCAGTCTACCTTGGGCACTGGTGGCTCTGAGTGTCTCGGTGCC 2750

Qy 2773 TGGGAGGGGG 2783

Db 2751 TGGGAGGGGG 2761

RESULT 11
 AX574592 2433 bp DNA linear PAT 07-JAN-2003
 LOCUS Sequence 3 from Patent WO0233087.
 DEFINITION AX574592
 ACCESSION AX574592
 VERSION AX574592.1 GI:27551850
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

1
 Edinger S., Gerlach, V., Macdonald, J.R., Malyankar, J.M.,
 Smithson, G., Millett, I., Peymann, J.A., Stone, D.J., Gunther, E.,
 Ellerman, K.J., Shmukets, R.A., Padigaru, M., Guo, X., Patturajan, M.,
 Tauripar, R.J., Burgess, C.E., Zerhusen, B.D., Kekuda, R., Spytek, K.A.,
 Gangolli, E.A., Fernandes, E.R. and Gorman, L.
 Proteins and nucleic acids encoding same
 Patent: WO 0233087-A 3 25-APR-2002;
 Curagen Corporation (US)

FEATURES
 Location/Qualifiers
 1..2433
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

BASE COUNT 567 a 656 c 674 g 536 t

ORIGIN
 Query Match 62.0%; Score 1795.6; DB 6; Length 2433;
 Best Local Similarity 82.0%; Pred. No. 0;
 Matches 2375; Conservative 0; Mismatches 34; Indels 489; Gaps 7;

Qy 1 CGCTCTCTGATGAAGCCCGCGCGGAGTGGCGGGCTTGGCGGCTGTGGGATGCTG 60

Db 22 CGCTCTCTGATGAAGCCCGCGCGGAGTGGCGGGCTTGGCGGCTGTGGGATGCTG 81

Qy 61 TTGGCGCAGTGGCGGAGCAGTGAAGTCCCGGGCGCTCCACACGCGGAAACCGCGG 120

Db 82 CTGGCGCAGTGGCGGAGCAGTGAAGTCCCGGGCGCTCCACACGCGGAAACCGCGG 141

Qy 121 TCGGACAGTGGGAGGAGTCCCGCGGCTCTCTCCCGGAGACCCCGCTCTCACC 180

Db 142 TCGGACAGTGGGAGGAGTCCCGCGGCTCTCTCCCGGAGACCCCGCTCTCACC 201

Qy 181 GCGATGTGCGCGCTGTTTTCCGAGGCACTGTGCGCATGGGACCCGAGCGGAGC 240

Db 202 GCGATGTGCGCGCTGTTTTCCGAGGCACTGTGCGCATGGGACCCGAGCGGAGC 261

Qy 241 GCTTGGGAGCGGAGTCCCGCGCTCTCTCCACCGCGGAGCGCGGCTGTGATGA 300

Db 262 GCTTGGGAGCGGAGTCCCGCGCTCTCTCCACCGCGGAGCGCGGCTGTGATGA 321

Qy 301 AAAGGCGGACATGGATGAAGCTGGAAACCATCTCTCAGCAAACTAAACAGAAACA 360

Db 322 AAAGGCGGACATGGATGAAGCTGGAAACCATCTCTCAGCAAACTAAACAGAAACA 381

Qy 361 GAAACCAAACTGTCATGTTCTCACTC--AATATGACCTGTCTCTGCTTACGAGTT 417

Db 382 GAAACCAAACTGTCATGTTCTCACTCCTGTAATATGACCTGTCTCTGCTTACGAGTT 441

Qy 418 GACCAAGGCGGATTAAGTGTCCCATGAAATCATGACCATCAGCGCGGAGAGAGCA 477

Db 442 GACCAAGGCGGATTAAGTGTCCCATGAAATCATGACCATCAGCGCGGAGAGAGCA 501

Qy 478 GTGGCGGTGTCCAGTGTGATCTTCACTTTCGGCTGAAAGGCCCGAGGACAGCTTC 537

Db 502 GTGGCGGTGTCCAGTGTGATCTTCACTTTCGGCTGAAAGGCCCGAGGACAGCTTC 561

Qy 538 CACATGGATCTGAGGACTTCCAGGAGCTAGTGGCTCTGCTTTATTTGTGAGAGCTTG 597

Db 562 CACATGGATCTGAGGACTTCCAGGAGCTAGTGGCTCTGCTTTATTTGTGAGAGCTTG 621

Qy 598 GGAAGAGAGGCACTTAAGTCTGTGAGACTTTTACCGCGAGAGACTTCTGTTTCTATCAA 657

Db 622 GGAAGAGAGGCACTTAAGTCTGTGAGACTTTTACCGCGAGAGACTTCTGTTTCTATCAA 681

Qy 658 GGCCTTTTGGATCACACAGAACTCCGCACTCGATGAGGAGAGTCTGTGAGGAGCTCC 717

Db 682 GGCCTTTTGGATCACACAGAACTCCGCACTCGATGAGGAGAGTCTGTGAGGAGCTCC 741

Qy 718 ACTCGCACTCTGAGCTCTGCAACAGTCAAGATGCTCCCGGAGAGTGTGACTTCGT 777

REFERENCE
AUTHORSTITLE
JOURNALFEATURES
source

RESULT 12
 AX574594
 LOCUS
 DEFINITION
 Sequence 5 from Patent W00233087.
 AX574594
 ACCESSION
 VERSION
 AX574594.1 GI:27551851
 KEYWORDS
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE
 AUTHORS
 Edinger S., Gerlach V., Macdougall J.R., Malyanar U.M.,
 Smithson G., Miller J., Peyman J.A., Stone D.J., Gunther E.,
 Ellerman K., Shinkels R.A., Padigaru M., Guo X., Paturajan M.,
 Taupier P.J., Burgess C.E., Zerkhusen B.D., Kekuda R., Szytek K.A.,
 Gangoli E.A., Fernandes E.R. and Gorman L.
 TITLE
 Protein and nucleic acids encoding same
 JOURNAL
 Patent: W0 0233087-A 5 25-APR-2002;
 Curren Corporation (US)
 FEATURES
 source
 Location/Qualifiers
 ..2902
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 BASE COUNT 699 a 752 c 777 g 674 t
 ORIGIN

Query Match 58.5%; Score 1695; DB 6; Length 2902;
 Best Local Similarity 78.1%; P=ed No. 0;
 Matches 2333; Conservative 0; Mismatches 315; Indels 339; Gaps 11;

QY 40 TTGGCGCGCTGTGGATGCTGTGTGCGCAGGTGGCGAGCAGTGGTCCGGGGGCTCC 99
 DB 1 TTGGCGCGCTGTGGATGCTGTGTGCGCAGGTGGCGAGCAGTGGTCCGGGGGCTCC 60
 QY 100 CACGAGCGCGAAACCGCGGTCGCGACAGCTGGAGGCGAGTCCCGCGGCTCTCTCC 159
 DB 61 CACGAGCGCGAAACCGCGGTCGCGACAGCTGGAGGCGAGTCCCGCGGCTCTCTCC 120
 QY 160 CGCGACCCCGCGCTCTCACCGGATGTGCGGCTGTGTTTCGCGAGCAGCTCGTGGCG 219
 DB 121 CGCGACCCCGCGCTCTCACCGGATGTGCGGCTGTGTTTCGCGAGCAGCTCGTGGCG 180
 QY 220 CATGGACCCCGCGAGCGCGCTGGAGCGCGAGGCTCGCGCTCTCTCTCCACCGCG 279
 DB 181 CATGGACCCCGCGAGCGCGCTGGAGCGCGAGGCTCGCGCTCTCTCTCCACCGCG 235
 QY 280 GGAGCGCGCGCTGTGATGGAACCGCGGACATGGATGAGTGGAAACATCGTTCT 339
 DB 236 ----- 235
 QY 340 CAGCAAACTAACACAGGACAGAAACCAACACTGCATGTCTCACTCAATATGACCTG 399
 DB 236 -----GACCTG 241
 QY 400 GTCTCTGCTTACAGGTTGACACAGGGCGATTACGTGCCATGAAATCATGCACCAT 459
 DB 242 GTCTCTGCTTACAGGTTGACACAGGGCGATTACGTGCCATGAAATCATGCACCAT 301
 QY 460 CAGCGGGGAGAGAGAGAGTGGCGCTGTGCGAGTTGAGTCTCTTCACTTCGGCTGAAA 519
 DB 302 CAGCGGGGAGAGAGAGAGTGGCGCTGTGCGAGTTGAGTCTCTTCCAGGTATGC 361
 QY 520 GCGCCACGAGCAGCTTCCACATGATCTGAGGACTCCAGGAGCTAGTGGCTCTGGC 579
 DB 362 AGAGCAGAGAGCTCAGACT-----GTGTGTGGAGGCTTTTCCATGCTAATTTCTCAGC 416
 QY 580 TTATTTGTCAGACCTTGGGAGAGAGAGGACGACTTAAGTCTGTGACAGCTTTACCGCAGAG 639
 DB 417 CGGGGTTTTTGAACCTTTTCCAATGTTGGCTCTCACTGGAGGAGACAGCATGCTTCCAAGA 476

640 GACTCTGTGTTCTATCAGGGCTCTTTGGGATCACAAGAAACTCGCCATCGCATGGAGG 699
 477 GAATAATAACAAATGCAA-----TGCTTGGAAGATCGGCTGCTTCAACCAAGAAAG 529
 700 AAGTCTGTGGAGGCTCCACATCGCACTCTGAAGCTCTGCAACAGTCAGTAAATGTCTCCCG 759
 530 TCTAATTTGTGTTTCTTTCTTTCTTTTATTTTCTTCTTCTTCTTCTTCTTCTTCTTCT 589
 760 GACAGTGTGACTTCTCGTCTGCTCAGTGTGCGGACACACAGCAGCAGATGTCAGAGG 819
 590 GAGCAGATTATTTCTTAAGGCCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 649
 820 CGGCACCTACAAGTGG-----AAGCCTTACACTCAAGTAGAAGCGGACTTATGCAAGCTTCTACT 877
 650 GCCCAAGGCGACTCGGCATCCAGCTCCAGCTGTACAGAGAGAGGCTCTGTGTGACTCTAAGG 709
 878 GTATGTCAGAGAGGATTTGATTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 937
 710 ACATGGGAGC-----TGSCACATCAACCCCTGCACAGCAGCA 747
 938 CATCTCGGAGGATAGCGTAACTTTGTGTATGATGGGATATGTGAGCTCAGTGTGTGT 997
 748 CTTTGGCTGGGATGCGACAAAGCAGCATTTCTGTGGAAGACGAGAAAT-----800
 998 CCACATCTCGGCACATGCCAGCTCCCAAGAGAGACCTCTTCTTCTTCTTCTTCTTCTTCTTCT 1057
 801 -----ACATGCCCCAGCTCCCAAGAGAGACCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 848
 1058 ATAGTCTTGTCTTACGGCATAGCGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1117
 849 ATAGTCTTGTCTTACGGCATAGCGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 908
 1118 AGCTGGAGACCTTGTGTGTGTCACAAAGAGATGATGCAAAACCATGCCATGAAATA 1177
 909 AGCTGGAGACCTTGTGTGTGTCACAAAGAGATGATGCAAAACCATGCCATGAAATA 968
 1178 TCACCACTTACGCTCAGTACTCAATGCTATCTGCTTTATTTCAAGA-----1229
 969 TCACCACTTACGCTCAGTACTCAATGCTATCTGCTTTATTTCAAGATGGAACAA 1028
 1230 -----1229
 1029 TAGGAGAAACATCAATTGCAATTTAGTCTGATTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1088
 1230 -----TG 1231
 1089 GACTGTGTATGATCACCAGCGACACCACTTAAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1148
 1232 GATTGTGGGAAAGATGGGACTGTGTCATGACCAAGCATCTTACTGCTGTCTGGATA 1291
 1149 GATTGTGGGAAAGATGGGACTGTGTCATGACCAAGCATCTTACTGCTGTCTGGATA 1208
 1292 TATGCTCTCGGAGAGACCTTGTGTCACACTTTTGGATTTGCCATTAAGTGGAAATGT 1351
 1209 TATGCTCTCGGAGAGATGAGCCCTGTGACACTTTTGGGATTTGCCATTAAGTGGAAATGT 1268
 1352 GTATGAAATATCGCAGCTCAGATTAATGAGATACAGCTTGTGGACTTGGGCTTCAACA 1411
 1269 GTATGAAATATCGCAGCTCAGATTAATGAGATACAGCTTGTGGACTTGGGCTTCAACA 1328
 1412 TTGCCCATCAGTCTGGACACACTTTGGCATGATCTTCTGATGGAGAGGACATGTGTA 1471
 1329 TTGCCCATCAGTCTGGACACACTTTGGCATGATCTTCTGATGGAGAGGACATGTGTA 1388
 1472 AAAAGTCCGAGGACACATCATGTCCCTTACATTTGGCAGAGCGCAATGAGCTTCTCT 1531
 1389 AAAAGTCCGAGGACACATCATGTCCCTTACATTTGGCAGAGCGCAATGAGCTTCTCT 1448
 1532 GGTCACTCTGAGCGCGAGTATCTACAAATTTCTTACAGCAGCGCTCAGCTATCTGCG 1591
 1449 GGTCACTCTGAGCGCGAGTATCTACAAATTTCTTACAGCAGCGCTCAGCTATCTGCG 1508
 1592 TTGCTGATCAGCCAAAGCCTGTGAGGAGATCAAGTATCTCTGAGAAATTTGCCAGAGAAAT 1651

Db	2553	GTGTTCCTGGGAATACCTCATGCTTCGTGTGGGACGACGAGACGACCCCTGCTCCACG	2613
Qy	2729	CGAGCTACACTTTTGGGCGCAATCGTGGCTCTGAGTCTTCCTCGTGTCTTCSGGAGGGGTAGCT	2788
Db	2613	CGAGCTACACTTTTGGGCGCAATCGTGGCTCTGAGTCTTCCTCGTGTCTTCSGGAGGGGTAGCT	2672
Qy	2789	GCCTTCCAGTCTGCTCTCTCGGACGACATGTGAGCTTTCAGCCACTTCGCTACATTGCAC	2848
Db	2673	GCCTTCCAGTCTGCTCTCTCGGACGACATGTGAGCTTTCAGCCACTTCGCTACATTGCAC	2732
Qy	2849	TGSCCTTTCTTGAATCCTTAATGAGCAGCCCGGGGCTTCCTCCCTGCCA	2895
Db	2733	TGSCCTTTCTTGAATCCTTAATGAGCAGCCCGGGGCTTCCTCCCTGCCA	2779

RESULT 13

BD170079

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BD170079

Novel ADAMTS family polypeptide and gene encoding the same.

BD170079

BD170079.1 GI:27875891

WO 0231163-A/1

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 4234)

(Obara, O., Nagase, T., Nomura, N., Yano, K., Murakami, K., Yasuda, S. and Kanazaki, K.

Novel ADAMTS family polypeptide and gene encoding the same

Patent: WO 0231163-A 1 18-APR-2002;

KAZUSA DNA RESEARCH INSTITUTE, MITSUBISHI PHARMA CORP., OSAMU OHARA, TAKAHIRO NAGASE, NOBUO NOMURA, KAZUHIRO YANO, KOJI MURAKAMI, SHINICHIRO YASUDA, KOJI KANZAKI

OS Homo sapiens (human)

PN WO 0231163-A/1

PD 11-APR-2002

PF 11-OCT-2001 WO 2001JP008913

PR 11-OCT-2000 JP 00P 311309, 02-APR-2001 JP 01P 102905 PI

OSAMU OHARA, TAKAHIRO NAGASE, NOBUO NOMURA, KAZUHIRO YANO, KOJI PI

MURAKAMI,

PI SHINICHIRO YASUDA, KOJI KANZAKI

PC C12N15/57, C12N9/64, C12M5/00, C12N1/19, C12N1/21, C12N1/15, C07K16/00

PC 40, C12Q1/68,

PC G01N33/50

CC Novel ADAMTS family polypeptide and gene encoding the same FH

Key Location/Qualifiers

FT CDS (3): 13088.

FEATURES

source

1..4234

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="GenBank:69606"

BASE COUNT 1036 a 1105 c 1130 g 965 t

ORIGIN

Query Match 56.5%; Score 1634.8; DB 6; Length 4234;

Best Local Similarity 93.3%; Pred. No. 0;

Matches 1772; Conservative 2; Indels 126; Gaps 1;

Qy 1010 ACATGCCCCAGCTCCCAAGGAAGACCTCTTCATCTTCCGAGTAGCTATAGTCTTGCT 10699

Db 157 ACATGCCCCAGCTCCCAAGGAAGACCTCTTCATCTTCCGAGTAGCTATAGTCTTGCT 216

Qy 1070 TACGCGATAGCGCTCTCTTCGAGGTCCCATAGAAATGAGAACTGAACGTGGAGACCT 1129

Db 217 TACGCGATAGCGCTCTCTTCGAGGTCCCATAGAAATGAGAACTGAACGTGGAGACCT 276

Qy 1130 TGGTGTGGTGGACAAAAGATGATCAAAACCATGGCCATGAAAATATCCACCACTTACG 11899

Db 277 TGGTGTGGTGGACAAAAGATGATCAAAACCATGGCCATGAAAATATCCACCACTTACG 336

Qy 1190 TGCTTCAGTACTCAACATGGTATCTGCTTTTATCAAGA----- 1229

697	CTGGACACACTTGTGGCATGATTTCAATGATGGAGAGGAGAAATGTCGAAAAAAGTCCGAGG	756
1484	GCAACATCATGTCCCTACATATTGGACAGAGCCAAATGGAGTCTTCTCTGTGTACACCTGCA	1543
757	GCAACATCATGTCCCTACATATTGGACAGAGCCAAATGGAGTCTTCTCTGTGTACACCTGCA	816
1544	GCGCCAGGTATCTACACAAATTTCTTAAAGCACCCGCTCAAGCTATCTTGCCTTGCTGATCAGC	1603
817	GCGCCAGGTATCTACACAAATTTCTTAAAGCACCCGCTCAAGCTATCTTGCCTTGCTGATCAGC	876
1604	CAAAAGCTGTGAAGGAATACAAAGTATCTTGAGAAATTCGCCAGGAGAAATATATGATGCAG	1663
877	CAAAAGCTGTGAAGGAATACAAAGTATCTTGAGAAATTCGCCAGGAGAAATATATGATGCAG	936
1664	ACACACAGTGCMAATGGCAGATTTCGGAGAGAAAGCCAAAGCTCTCTCATGCTGGAGCTTAAAA	1723
937	ACACACAGTGCMAATGGCAGATTTCGGAGAGAAAGCCAAAGCTCTCTCATGCTGGAGCTTAAAA	996
1724	AGGACATCTGTAAAGCCCTCTGGTGCACATCGTATGGAAGAAATGTGAGACTTAAATTTTA	1783
997	AGGACATCTGTAAAGCCCTCTGGTGCACATCGTATGGAAGAAATGTGAGACTTAAATTTTA	1056
1784	TGCCACGACGAGAGSCACAAATTTGTGGGCATGACATGTGCTGCCGGGAGAGCAGTGTG	1843
1057	TGCCACGACGAGAGSCACAAATTTGTGGGCATGACATGTGCTGCCGGGAGAGCAGTGTG	1116
1844	TGAAATATGTTGATGAAGGCCCCACAGCCACCAATGCGCACTGGTCGGAGCTGTCTTCTT	1903
1117	TGAAATATGTTGATGAAGGCCCCACAGCCACCAATGCGCACTGGTCGGAGCTGTCTTCTT	1176
1904	GGTCCCATGTCTCCAGCAGCTGCGAGAGGGAGATCTCATGAGAGTGGCTCTGCACCA	1963
1177	GGTCCCATGTCTCCAGCAGCTGCGAGAGGGAGATCTCATGAGAGTGGCTCTGCACCA	1236
1964	ACCCGACGACATCCGATCGAGAGAGTCTCTGTAGAGGCTCCACATCGCATCTGAGACTCT	2023
1237	ACCCGACGACATCCGATCGAGAGAGTCTCTGTAGAGGCTCCACATCGCATCTGAGACTCT	1296
2024	GCACAGCTCAGAAATATCCCGCGACAGTGTGACTTCCGTGCTGCTCAAGTGTGCCGAGC	2083
1297	GCACAGCTCAGAAATATCCCGCGACAGTGTGACTTCCGTGCTGCTCAAGTGTGCCGAGC	1356
2084	ACACACGACGAGATTCAGAGGGCGGCATCAAGTGGAGGCTTACATCAAGTGAAGAG	2143
1357	ACACACGACGAGATTCAGAGGGCGGCATCAAGTGGAGGCTTACATCAAGTGAAGAG	1416
2144	ATCAGGACTTATGCAAACTCTACTGTTATCGGAGAGGATTTGATTTCTCTTTCTTTGT	2203
1417	ATCAGGACTTATGCAAACTCTACTGTTATCGGAGAGGATTTGATTTCTCTTTCTTTGT	1476
2204	CAATTAAGTGCMAAGATGGGACTCCATGCTCGGAGGATACCGTAATGTTGTTATAGATG	2263
1477	CAATTAAGTGCMAAGATGGGACTCCATGCTCGGAGGATACCGTAATGTTGTTATAGATG	1536
2264	GGATATGTGAGAGATTTGGATGCTGCACAAATGTCCTTGGACTGATGCTGTTCGAGAGTCT	2323
1537	GGATATGTGAGAGATTTGGATGCTGCACAAATGTCCTTGGACTGATGCTGTTCGAGAGTCT	1596
2324	GTGGGGTGTGPAACGGGATATCTCAGGCTGCGAGATTCACAGGGGTCTCTACACCAAGC	2383
1597	GTGGGGTGTGPAACGGGATATCTCAGGCTGCGAGATTCACAGGGGTCTCTACACCAAGC	1656
2384	ACACACACGCAACAGATATATCAATGGTGCACATTCCTTGGGAGCCCGGAGATATCC	2443
1657	ACACACACGCAACAGATATATCAATGGTGCACATTCCTTGGGAGCCCGGAGATATCC	1716
2444	GGATCTATGAATGAGCGTCTCTACCTCTCATATTTCTGTGGGCAATGCCCTGAGAGGT	2503
1717	GGATCTATGAATGAGCGTCTCTACCTCTCATATTTCTGTGGGCAATGCCCTGAGAGGT	1776
2504	ACTTACCTGAATGGGCACTGGAACCTTGACATGGCCCGGCCGGGTACAAATTTTCGGGCACTA	2563
1777	ACTTACCTGAATGGGCACTGGAACCTTGACATGGCCCGGCCGGGTACAAATTTTCGGGCACTA	1836

[illegible]

```
|||||
Db 5306 AGATCAGGACTTATGCAAACTCTACTGTATCGCAGAGGATTTGATTTCTTCTTTT 5365
QY 2202 GTCAAAATAAGTCAAGATGGGACTCCATGCTCCGGAGGATAGCCGTAAATGTTGTATAGA 2261
Db 5366 GTCAAAATAAGTCAAGATGGGACTCCATGCTCCGGAGGATAGCCGTAAATGTTGTATAGA 5425
QY 2262 TGGGATATGTGAGAGAGTGTGGATGTGACAAATGTCTTTGGATCTGATCTGTTTGAAGACGT 2321
Db 5426 TGGGATATGTGAGAGAGTGTGGATGTGACAAATGTCTTTGGATCTGATCTGTTTGAAGACGT 5485
QY 2322 CTGTGGGGTGTGTAAACGGGAATAACTCAGCCCTGCACCAATTCACAGGGTCTCTACACCA 2381
Db 5486 CTGTGGGGTGTGTAAACGGGAATAACTCAGCCCTGCACCAATTCACAGGGTCTCTACACCA 5545
QY 2382 GCACCAACCAACCAACCAAGTATTATCAATGTGTACCAATTCCTTCTGGAGCCCGAGTAT 2441
Db 5546 GCACCAACCAACCAACCAAGTATTATCAATGTGTACCAATTCCTTCTGGAGCCCGAGTAT 5605
QY 2442 CCGCATCTATGAAATGAAAGTCTCTTACCTCTCAATTTCTGTGGCCAAATGCCCTCAGAAG 2501
Db 5606 CCGCATCTATGAAATGAAAGTCTCTTACCTCTCAATTTCTGTGGCCAAATGCCCTCAGAAG 5665
QY 2502 GTACTACTGAAATGGGCACTGGAACCGTGGACTGGCCCGCGGTACAAATTTTGGGCAAC 2561
Db 5666 GTACTACTGAAATGGGCACTGGAACCGTGGACTGGCCCGCGGTACAAATTTTGGGCAAC 5725
QY 2562 TACTTTCGACTACAGACGGTCTCTAATATGAGCCCGAGACTTAATTCGCTACTGGACCAAC 2621
Db 5726 TACTTTCGACTACAGACGGTCTCTAATATGAGCCCGAGACTTAATTCGCTACTGGACCAAC 5785
QY 2622 CAACGACACACTGATTTGTGAGCTGCTGTTTCAGGGAAGGAACCCGGGTGTTGCCCTGGGA 2681
Db 5786 CAACGACACACTGATTTGTGAGCTGCTGTTTCAGGGAAGGAACCCGGGTGTTGCCCTGGGA 5845
QY 2682 ATACTCCATGCTCTGCTTGGGGAACCGGAGAGAGAGCTTCTGCTCCAGCCCAAGCTACACTTG 2741
Db 5846 ATACTCCATGCTCTGCTTGGGGAACCGGAGAGAGAGCTTCTGCTCCAGCCCAAGCTACACTTG 5905
QY 2742 GGCCATCGCGCTCTGAGTGTCTCGTCTCTCGGAGCGGG 2783
Db 5906 GGCCATCGCGCTCTGAGTGTCTCGTCTCTCGGAGCGGG 5947
```

Search completed: October 28, 2003, 21:53:52
Job time : 10283 secs

Result No.	Score	%		Query Match	Length	DB	ID	Description
1	2895	100.0	2895	24	ABSS5936			Human ADAM-TS 7-li
2	2675.4	92.4	2997	24	ABSS5933			Human zinc metallo
3	1877.8	64.9	3675	24	AAS371174			Human metalloprote
4	1877.8	64.9	5610	24	ABK4982			Human CDNA encodin
5	1877.8	64.9	8435	24	ABK49825			plasmid pFastBac1-
6	1877.8	64.9	8505	24	ABK49826			plasmid pFastBac1-
7	1874.6	64.8	3675	24	AAL43654			Human MDTF9 protea
8	1874.6	64.8	3675	25	ABSS57767			Novel human protea

KW congenital diarrhoea; respiratory disease; gastro-intestinal disease;
 KW muscle disorder; bone disorder; joint disorder; skeletal disorder;
 KW haematopoietic disorder; urinary system disorder; osteoporosis; ds;
 KW dental disease; dental infection; growth disorder; reproductive disorder;
 KW hypogonadism; fertility disorder; viral infection; bacterial infection;
 KW parasitic infection; metabolic pathway modulation; gene therapy; gene;
 KW zinc metalloproteinase; ADM-TS 7; alpha-2-macroglobulin precursor;
 KW ileal sodium/bile acid cotransporter; prohibitin; MT; C194; spinesin;
 KW macrophage stimulating protein precursor; fatty acid-binding protein;
 KW gap junction beta-5 protein; hepsin/plasma transmembrane serine protease;
 KW single nucleotide polymorphism; SNP.
 XX Homo sapiens.
 OS
 FH Key Location/Qualifiers
 FT variation replace (472,G)
 FT /*tag= a
 FT /*standard_name= "Single nucleotide polymorphism (SNP)"
 FT variation replace (481,C)
 FT /*tag= b
 FT /*standard_name= "Single nucleotide polymorphism (SNP)"
 FT variation replace (1121,C)
 FT /*tag= c
 FT /*standard_name= "Single nucleotide polymorphism (SNP)"
 FT variation replace (1516,C)
 FT /*tag= d
 FT /*standard_name= "Single nucleotide polymorphism (SNP)"
 FT variation replace (1566,C)
 FT /*tag= e
 FT /*standard_name= "Single nucleotide polymorphism (SNP)"
 XX WO200233087-A2.
 XX
 XX 25-APR-2002.
 XX
 XX 17-OCT-2001; 2001WO-US32496.
 XX
 XX 17-OCT-2000; 2000US-241040P.
 XX 17-OCT-2000; 2000US-241058P.
 XX 17-OCT-2000; 2000US-241063P.
 XX 17-OCT-2000; 2000US-241063P.
 XX 20-OCT-2000; 2000US-242153P.
 XX 23-OCT-2000; 2000US-242182P.
 XX 23-OCT-2000; 2000US-242182P.
 XX 23-OCT-2000; 2000US-242111P.
 XX 23-OCT-2000; 2000US-242117P.
 XX 24-OCT-2000; 2000US-242860P.
 XX 24-OCT-2000; 2000US-242861P.
 XX 29-DEC-2000; 2000US-252639P.
 XX 20-FEB-2001; 2001US-262913P.
 XX 25-APR-2001; 2001US-286324P.
 XX 29-MAY-2001; 2001US-294109P.
 XX 09-JUL-2001; 2001US-303698P.
 XX 16-OCT-2001; 2001US-0981151.
 XX
 XX (CURA-) CURAGEN CORP.
 XX
 XX Edinger S, Gerlach V, MacDougall JP, Malyankar UM, Smithson G;
 XX Millet I, Peyman JA, Stone DJ, Gunther E, Ellerman K, Shinkets RA;
 XX Padigara M, Guo X, Patterson M, Taupier R, Burgess CE;
 XX Zethusen BD, Kekula R, Spytek KA, Gangolli EA, Fernandes ER;
 XX Gorman L;
 XX
 XX WPI: 2002-590434/63.
 XX P-PSDB: ABG76894.
 XX
 XX Cytoplasmic, nuclear, membrane bound and secreted polypeptides and
 XX nucleic acids encoding the polypeptides for diagnosing and treating
 XX e.g. cancer, Alzheimer's disease, cardiomyopathy, metabolic disease and
 XX diabetes.
 XX
 XX Claim 8; Page 12; 305pp; English.
 XX
 XX The present invention relates to new NOVX (NOV1-10) polypeptides. The

CC molecules of the invention are useful for treating or preventing a
 CC NOVX-associated disorder, such as cardiomyopathy, atherosclerosis, or
 CC a disorder related to cell signal processing and metabolic pathway
 CC modulation in humans. NOVX polypeptides, nucleic acids and antibodies
 CC are useful for treating or preventing disorders or syndromes including
 CC breast cancer, Alzheimer's disease, epilepsy, Huntington's disease,
 CC anxiety, behavioural disorders, multiple sclerosis, myasthenia gravis,
 CC neurodegeneration, Parkinson's disease, pain, stroke, autoimmune
 CC disease, allergies, addiction, asthma, endometriosis, graft versus host
 CC disease, systemic lupus erythematosus, scleroderma, transplantation,
 CC psoriasis, Crohn's disease, HIV (human immunodeficiency virus) infection,
 CC atherosclerosis, cirrhosis, rheumatoid arthritis, diabetes,
 CC thrombocytopenia, bleeding disorders, metabolic disorders, obesity,
 CC glucose transport defect, glomerulonephritis, hypercalcaemia, polycystic
 CC kidney disease, pancreatitis, renal tubular acidosis, skin disorders,
 CC congenital diarrhoea, respiratory disease, gastro-intestinal diseases,
 CC muscle, bone, joint and skeletal disorders, haematopoietic disorders,
 CC urinary system disorders, osteoporosis, dental disease and infection,
 CC growth and reproductive disorders, hypogonadism, fertility, and/or other
 CC pathologies and disorders, viral, bacterial, or parasitic infections.
 CC The present nucleic acid sequence encodes a NOVX protein of the

XX SQ Sequence 2997 BP; 728 A; 783 C; 809 G; 673 T; 4 other:

Query Match 92.4%; Score 2675.4; DB 24; Length 2997;
 Best Local Similarity 94.9%; Pred. No. 0;
 Matches 2866; Conservative 0; Mismatches 5; Indels 150; Gaps 4;

OY 1 CGCTCTTGATCAAGCCCGCGCGGATGGCGGGCTTGCGGGGCTGTGGATGCTG 60
 DB 1 CGCTCTTGATCAAGCCCGCGCGGATGGCGGGCTTGCGGGGCTGTGGATGCTG 60
 OY 61 TTGCGCGAGGTGGCGGAGGTGAGTGGCGGGCTCCACAGCGGAAACCGCGGG 120
 DB 61 TTGCGCGAGGTGGCGGAGGTGAGTGGCGGGCTCCACAGCGGAAACCGCGGG 120
 OY 121 TCAGGACAGCTGGAGGCGAGTCCCGCGGGCTCTCTCCCGGGACCCCGCGCTCACC 180
 DB 121 TCAGGACAGCTGGAGGCGAGTCCCGCGGGCTCTCTCCCGGGACCCCGCGCTCACC 180
 OY 181 GCGATGTCGCGCTGTGTTTCCGAGGACCTCGTGGCGGATGGGACCCGACGCGGAGC 240
 DB 181 GCGATGTCGCGCTGTGTTTCCGAGGACCTCGTGGCGGATGGGACCCGACGCGGAGC 240
 OY 241 GCCTGGGAGCCCGAGCGTCCCGCGTCTCTCCACCGCGGAGCGCGGGCTGGATGGA 300
 DB 241 GCCTGGGAGCCCGAGCGTCCCGCGTCTCTCCACCGCGGAGCGCGGGCTGGATGGA 300
 OY 301 AAAGGCGGAGCATGATGAAAGTGGAAACCATGTTCTACGAAATACACAGGAACA 360
 DB 301 AAAGGCGGAGCATGATGAAAGTGGAAACCATGTTCTACGAAATACACAGGAACA 360
 OY 361 GAAACCAAAACCTGCATGTTCTCACTCAATATGACCTGGTCTCTGCTACGAGTTGAC 420
 DB 361 GAAACCAAAACCTGCATGTTCTCACTCAATATGACCTGGTCTCTGCTACGAGTTGAC 420
 OY 421 CACAGGCGGATTAAGTGTCCATGAAATCATGCAACCATCAGCGCGGAGAGAGCAGTG 480
 DB 421 CACAGGCGGATTAAGTGTCCATGAAATCATGCAACCATCAGCGCGGAGAGAGCAGTG 480
 OY 481 GCGTGTGCGAGGTGAGTCTCTTCACTTGGTGAAGGCCCCAGGACGACCTTCCAC 540
 DB 481 GCGTGTGCGAGGTGAGTCTCTTCACTTGGTGAAGGCCCCAGGACGACCTTCCAC 540
 OY 541 ATGATCTCAGGACTTCCAGCAGCTAGTGGCTCTGGCTCTATTGTGAGACGCTGGGA 600
 DB 541 ATGATCTCAGGACTTCCAGCAGCTAGTGGCTCTGGCTCTATTGTGAGACGCTGGGA 600
 OY 601 AAGCAGGCACTAAGTCTGTGAGACTTTACCGCCAGAGGACTTCTGTTTATCAAGGC 660
 DB 601 AAGCAGGCACTAAGTCTGTGAGACTTTACCGCCAGAGGACTTCTGTTTATCAAGGC 660

Qy	661	TCCTTGGCATCACAAGAACTCCGCATGTGCAATGGAGGAGATCTCTGCGAGGCTCCACT	720
Dy	661	TCCTTGGCATCACAAGAACTCCGCATGCATATGAGGAGATCTCTGCGAGGCTCCACT	720
Qy	721	CGCACTCTGAGCTCTGCAACGATCTCCCGGAGACAGTGTGCACTTCCCGTCT	780
Dy	721	CGCACTCTGAGCTCTGCAACGATCTCCCGGAGACAGTGTGCACTTCCCGTCT	780
Qy	781	GCTCAGTGTGCCGAGCACCAACGACGATTTAGAGGGGCGCACTACAAGTGGAGACCT	840
Dy	781	GCTCAGTGTGCCGAGCACCAACGACGATTTAGAGGGGCGCACTACAAGTGGAGACCT	840
Qy	841	TACACTCAAGTAGAAGCGCACTTATCGAAACTCTACTGTATCGCAGAAAGATTGATTTTC	900
Dy	841	TACACTCAAGTAGAAGCGCACTTATCGAAACTCTACTGTATCGCAGAAAGATTGATTTTC	900
Qy	901	TTCTTTTCTTCTCAAAATAAGTCMAAGATGGGATCCCAATGCTGGAGGATAGCGGTAAAT	960
Dy	901	TTCTTTTCTTCTCAAAATAAGTCMAAGATGGGATCCCAATGCTGGAGGATAGCGGTAAAT	960
Qy	961	GTTTGTATAGATGGGATATGTGAGCTCAGTGTGTGTGCCAATCTGTGGCACAATGCCCCAG	1020
Dy	961	GTTTGTATAGATGGGATATGTGAGCTCAGTGTGTGTGCCAATCTGTGGCACAATGCCCCAG	1020
Qy	1021	CTCTCCAAAGGAAGACCTTCTCATCTTCCAGATGATGATATAAGTCTTGCTTAGCGCATAAAG	1080
Dy	1021	CTCTCCAAAGGAAGACCTTCTCATCTTCCAGATGATGATATAAGTCTTGCTTAGCGCATAAAG	1080
Qy	1081	CGCTCTCTTCTGAGGTCCTCATAGAAATGAGAAATCGAAAGCTGGAGACCTTGTGTGGTCT	1140
Dy	1081	CGCTCTCTTCTGAGGTCCTCATAGAAATGAGAAATCGAAAGCTGGAGACCTTGTGTGGTCT	1140
Qy	1141	GCACAAAGATATGACAAAACCATGCGCATGAAAATATCAACCACTACTGCTCAGCATAT	1200
Dy	1141	GCACAAAGATATGACAAAACCATGCGCATGAAAATATCAACCACTACTGCTCAGCATAT	1200
Qy	1201	CTCAACATGATATCTGCTTTATTCAAAGA-----	1259
Dy	1201	CTCAACATGATATCTGCTTTATTCAAAGAATGGAACAATAGAGGAAACATCAACATTGCA	1260
Qy	1220	-----	1229
Dy	1261	ATTGTAGGTCGATCTTCTTAGAGATGGAACGACGAGACCTGGTATGATCAACACGCA	1320
Qy	1321	-----TGAGTGAATGGGGAAGATGCGAT	1254
Dy	1321	GACCAACCTTAAGTAGCTCTGCGAGTGGCAGTCTGGATTCATGGGGAAGATGGAGCT	1380
Qy	1255	CGTATGACGACGATCTTACATGACTGCTGGTATATGTTCTCTGGAGAAATGACGCC	1314
Dy	1381	CGTATGACGACGATCTTACATGACTGCTGGTATATGTTCTCTGGAGAAATGACGCC	1440
Qy	1315	TCGTGACATTTGGGATTCACCATAGATGGGAATGTAGTATAATATCGCAGCTGCAGC	1374
Dy	1441	TCGTGACATTTGGGATTCACCATAGATGGGAATGTAGTATAATATCGCAGCTGCAGC	1500
Qy	1375	ATTATAGAGATACAGCTCTTGACCTGGGCTTCAACATTGCCCATGAGTCTGGACACAAC	1434
Dy	1501	ATTATAGAGATACAGCTCTTGACCTGGGCTTCAACATTGCCCATGAGTCTGGACACAAC	1560
Qy	1435	TTTGGCATGATTCATGATGAGGAAGGCAACATGTGTAAAAGATCCGAGGGGCAACATCATG	1494
Dy	1561	TTTGGCATGATTCATGATGAGGAAGGCAACATGTGTAAAAGATCCGAGGGGCAACATCATG	1620
Qy	1495	TCCCTCATTTGGCAGAGACGCAATGAGTCTTCTCTGCTGTACCTCTCGAGCGCCAGTAT	1554
Dy	1621	TCCCTCATTTGGCAGAGACGCAATGAGTCTTCTCTGCTGTACCTCTCGAGCGCCAGTAT	1680
Qy	1555	CTACACAAATTTCTAAGCACCGCTCAAGCTATCTGCCTTGTGTATCAGCAAGAGCCTGTG	1614
Dy	1681	CTACACAAATTTCTAAGCACCGCTCAAGCTATCTGCCTTGTGTATCAGCAAGAGCCTGTG	1740
Qy	1615	AAGGAATACAAAGTATCTCTGTGAAAATTTGCCAGGAGGAATATATATGATGCAACACAGATGC	1674

Db	1741	AAAGGATACAGTATCTTGAGAAATTCGCCAGGAAATATATGATCAACACACAGATGC	1800
Qy	1675	AAATGGCAGTTCGGAGAGAAAGCAAGCTCTGCATCTCGACTTTTAAAGAGCACTCTGT	1734
Db	1801	AAATGGCAGTTCGGAGAGAAAGCAAGCTCTGCATCTCGACTTTTAAAGAGCACTCTGT	1860
Qy	1735	AAAGCCCTGTGTGCCATCTGTTTGAAGAGAAATGTGAGACTTAAATTTATGCGACAGCA	1794
Db	1861	AAAGCCCTGTGTGTGCATCTGTTTGAAGAGAAATGTGAGACTTAAATTTATGCGACAGCA	1920
Qy	1795	GAAGGCACAAATTTGTGGCATGACATGTGTGTGCCGGGAGAGACAGTGTGTGAATATGTG	1854
Db	1921	GAAGGCACAAATTTGTGGCATGACATGTGTGTGCCGGGAGAGACAGTGTGTGAATATGTG	1980
Qy	1855	GATGAAGGCCCAAGCCCAAGCCATGAGCCATCTGTCGACATGTGTCTCTTTGGTCCCATGC	1914
Db	1981	GATGAAGGCCCAAGCCCAAGCCATGAGCCATCTGTCGACATGTGTCTCTTTGGTCCCATGC	2040
Qy	1915	TCCAGGACTCTGGAGAGGGGGAGTATCTCATAGAGTGCCTCTGCACCAACCCCAAGCCA	1974
Db	2041	TCCAGGACTCTGGAGAGGGGGAGTATCTCATAGAGTGCCTCTGCACCAACCCCAATCCA	2100
Qy	1975	TCCGATCGAGAGGAAGTTCTGTGAGGGCTCACTTGCACACTCTGAAGTCTTGCACACAGTCAG	2034
Db	2101	TCCGATCGAGAGGAAGTTCTGTGAGGGCTCACTTGCACACTCTGAAGTCTTGCACACAGTCAG	2160
Qy	2035	AAATGTCCCGGGACAGTGTGACTTCGGTGTCTCACTGTGCCGAGCAGACACAGCA	2094
Db	2161	AAATGTCCCGGGACAGTGTGACTTCGGTGTCTCACTGTGCCGAGCAGACACAGCA	2220
Qy	2095	CGATTCAGAGGGCGGCACTACAGGTGGAAGCTTACACTCAAGTAGAGATCAGGACTTA	2154
Db	2221	CGATTCAGAGGGCGGCACTACAGGTGGAAGCTTACACTCAAGTAGAGATCAGGACTTA	2262
Qy	2155	TGCAAACTCTACTGTATGTCGAGAGGATTTGATTTCTTTCTTTGTGCAAAATAAGTC	2214
Db	2263	TGCAAACTCTACTGTATGTCGAGAGGATTTGATTTCTTTCTTTGTGCAAAATAAGTC	2322
Qy	2215	AAAGTGGGACTCCACTCTGGAGGATAGCGGTATGTGTTGTATGATGGGATGTGAG	2274
Db	2323	AAAGTGGGACTCCACTCTGGAGGATAGCGGTATGTGTTGTATGATGGGATGTGAG	2382
Qy	2275	AGAGTGGATGTGACAAATGCTCTGGATCTGATCTGCTGTGAAGACTCTGTGGGGTGTG	2334
Db	2383	---NTTGATGTGACAAATGCTCTGGATCTGATCTGCTGTGAAGACTCTGTGGGGTGTG	2439
Qy	2335	AAAGCGAATACCTGAGCTGGAGCATCTACAGGGTCTCTACACAGCAGCCACACACC	2394
Db	2440	AAAGCGAATACCTGAGCTGGAGCATCTACAGGGTCTCTACACAGCAGCCACACACC	2499
Qy	2395	AACCAAGTATATACATGCTGTACCATTCCTCTCGAGCCCGGGGTATCGGCATCTATGA	2454
Db	2500	AACC---ATATACATGCTGTACCATTCCTCTCGAGCCCGGGGTATCGGCATCTATGA	2556
Qy	2455	ATGAACCTCTACTCCTCTACATTTCTGTGCGCAATGGCCCTCAGAAGGTACTACTGGAAT	2514
Db	2557	ATGAACCTCTACTCCTCTACATTTCTGTGCGCAATGGCCCTCAGAAGGTACTACTGGAAT	2616
Qy	2515	GGGCACTGGAGCGTGGACTGGCCCGCGCGGTACAAATTTTGGGCACTACTTTTCGACTAC	2574
Db	2617	GGGCACTGGAGCGTGGACTGGCCCGCGCGGTACAAATTTTGGGCACTACTTTTCGACTAC	2676
Qy	2575	AGAAGGCTCTATATATGAGCCCGAGAACTTATTCGCTACTTGGACCAACCAAGAGACACTG	2634
Db	2677	AGAAGGCTCTATATATGAGCCCGAGAACTTATTCGCTACTTGGACCAACCAAGAGACACTG	2736
Qy	2635	ATTGTGAGCTCTGTTTTCAGGAGAGAAACCGGGGTGTGCTCGGAATACTCTCATGCGCT	2694
Db	2737	ATTGTGAGCTCTGTTTTCAGGAGAGAAACCGGGGTGTGCTCGGAATACTCTCATGCGCT	2796
Qy	2695	CGCTTGGGACACGAGAGACGCGCCCTTGGCCGAGCCGAGCTACACTTGGGCCCATCTGGTGGC	2754

QY	1159	AACTGCGCCATGAAATATCAACCACTACGTCTCAAGATCTCAACATGATCTGCT	1218
DB	913	AACTGCGCCATGAAATATCAACCACTACGTCTCAAGATCTCAACATGATCTGCT	972
QY	1219	TTATTCAAGA-----	1229
DB	973	TTATTCAAGA-----	1032
QY	1230	-----	1229
DB	1033	CTGAAGATGAACAGCCAGGAGCTGGTATAGTCAACGAGACACACCTTAAATGAC	1092
QY	1230	-----TGAGTATGAGGAAAGATGGAGCTCTCATGACACAGCCATC	1272
DB	1093	TTTGGCAGTGGAGCTGGATATGATGAGGAAAGATGGAGCTCTCATGACACAGCCATC	1152
QY	1273	TTATGACTGGCTGGATATATGTTCTGGAAAGATGAGCTCTGTGAACCTTTGGATTT	1332
DB	1153	TTATGACTGGCTGGATATATGTTCTGGAAAGATGAGCTCTGTGAACCTTTGGATTT	1212
QY	1333	GCACCAATAGTGAATATGATGATTAATATGAGCTGACCAATTAATGAGATACAGGT	1392
DB	1213	GCACCAATAGTGAATATGATGATTAATATGAGCTGACCAATTAATGAGATACAGGT	1272
QY	1393	TTTGGACTGGCTTCAACCAATGAGCTGAGCTGACCACTTTGGAGATATCATAT	1452
DB	1273	TTTGGACTGGCTTCAACCAATGAGCTGAGCTGACCACTTTGGAGATATCATAT	1332
QY	1453	GGAGAGGAGACATGTTGAAAGTTCGGAGGAGCAATCATGCTCCATCACTGAGCA	1512
DB	1333	GGAGAGGAGACATGTTGAAAGTTCGGAGGAGCAATCATGCTCCATCACTGAGCA	1392
QY	1513	GGATAGGAGCTTCTCTGGTGGCTGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT	1572
DB	1393	GGATAGGAGCTTCTCTGGTGGCTGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT	1452
QY	1573	ACCCCTCAAGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1632
DB	1453	ACCCCTCAAGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1512
QY	1633	GAGAAATTCGAGGAGAAATATATGATGCAACACACACACTGCTGAGGAGTACAGTAT	1692
DB	1513	GAGAAATTCGAGGAGAAATATATGATGCAACACACACACTGCTGAGGAGTACAGTAT	1572
QY	1693	AAAGCCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1752
DB	1573	AAAGCCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1632
QY	1753	CGTATGAGGAGAAATGAGACTTAATTTATGCTGAGGAGGAGGAGGAGGAGGAGGAG	1812
DB	1633	CGTATGAGGAGAAATGAGACTTAATTTATGCTGAGGAGGAGGAGGAGGAGGAGGAG	1692
QY	1813	CATGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1872
DB	1693	CATGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1752
QY	1873	ACCCATGGCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1932
DB	1753	ACCCATGGCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1812
QY	1933	GGAGTATCTCATGAGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1992
DB	1813	GGAGTATCTCATGAGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1872
QY	1993	TGTGAGGGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2052
DB	1873	TGTGAGGGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1932
QY	2053	GTTGACTTCCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2112
DB	1933	GTTGACTTCCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1992
QY	2113	TACAAGTGGAGAGCTTACACTCAAGTAGAAGATCAGGACTTATGACAACTCTACTGTATC	2172
DB	1993	TACAAGTGGAGAGCTTACACTCAAGTAGAAGATCAGGACTTATGACAACTCTACTGTATC	2052
QY	2173	GCAGAGAGATTTGATTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT	2232
DB	2053	GCAGAGAGATTTGATTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT	2112
QY	2233	TCGAGGATAGACCGGTAAATGTTTATAGATGGATATGTGAGAGAGTGGATGTGACAAAT	2292
DB	2113	TCGAGGATAGACCGGTAAATGTTTATAGATGGATATGTGAGAGAGTGGATGTGACAAAT	2172
QY	2293	GTCCTTGGATGTGATGCTGTTGAAAGAAGTCTGTTGAGAGAGTCTGTTGAGAGAGTCT	2352
DB	2173	GTCCTTGGATGTGATGCTGTTGAAAGAAGTCTGTTGAGAGAGTCTGTTGAGAGAGTCT	2232
QY	2353	TCGAGATTCACAGAGGTCTCTACACCAAGGACCAACCAACCAACCAACCAACCAACCA	2412
DB	2233	TCGAGATTCACAGAGGTCTCTACACCAAGGACCAACCAACCAACCAACCAACCAACCA	2292
QY	2413	GTCACCATTCCTTCTGAGGCGGAGATTCGCGATCTATGAAATGAAAGTCTCTACCTCC	2472
DB	2293	GTCACCATTCCTTCTGAGGCGGAGATTCGCGATCTATGAAATGAAAGTCTCTACCTCC	2352
QY	2473	TACATTTCTGTCGGAATGCGCTTGAAGAGTCTACCTGAGTGGGACCTGGACCTGGAC	2532
DB	2353	TACATTTCTGTCGGAATGCGCTTGAAGAGTCTACCTGAGTGGGACCTGGACCTGGAC	2412
QY	2533	TGGGCGGCGGTTACAAATTTTGGGCGACTACTTTGGACTACAGACGCTCTATATGAG	2592
DB	2413	TGGGCGGCGGTTACAAATTTTGGGCGACTACTTTGGACTACAGACGCTCTATATGAG	2472
QY	2593	CCGAGAACTTAAATGCTGCTGAGGCGGAGCCAAACAGAGACCTGATTTGAGGCTGCT	2652
DB	2473	CCGAGAACTTAAATGCTGCTGAGGCGGAGCCAAACAGAGACCTGATTTGAGGCTGCT	2532
QY	2653	CHGSGAGAGACACCGGGTGTTCCTGGGAAATACCTCATGCTGCTTGGGAGCGAGAG	2712
DB	2533	CHGSGAGAGACACCGGGTGTTCCTGGGAAATACCTCATGCTGCTTGGGAGCGAGAG	2592
QY	2713	CHGCCCCCTGCCCCAGGCTACACTTGGGCGCACTGCTGCTGCTGCTGCTGCTGCTGCT	2772
DB	2593	CHGCCCCCTGCCCCAGGCTACACTTGGGCGCACTGCTGCTGCTGCTGCTGCTGCTGCT	2652
QY	2773	TGCGGAGGAGG 2783	
DB	2653	TGCGGAGGAGG 2663	
RESULT 4			
ABK49822			
ID	ABK49822	standard; CDNA; 5610 BP.	
XX	ABK49822;		
AC	ABK49822;		
XX	15-JUL-2002 (first entry)		
DT	15-JUL-2002 (first entry)		
XX	Human cDNA encoding ADAMTS protein #1.		
DE	Human; ss; gene; ADAMTS; cytotatic; antidiabetic; antirheumatic; SNP;		
XX	antiarthritic; antitumor; vulvar; neovascularisation; angiodysplasia;		
KW	diabetic omentopathy; chronic rheumatoid arthritis; gene therapy;		
KW	refractory skin ulcer; gastric ulcer; post-operative healing failure;		
KW	repolyspondin type 1 domain; sexual cycle; tumour; sp-syndrome deletion;		
KW	chromosome 5p15.2-15.3; Cxi-da-chat syndrome;		
KW	single nucleotide polymorphism.		
OS	Homo sapiens.		
XX	Key	Location/Qualifiers	
XX	770..4444		
XX	/*tag= a		
FT			

FT variation /product= "ADAMTS"
 FT replace (821,T)
 FT /standard_name= "Single nucleotide polymorphism"
 FT replace (1079,T)
 FT /tag= c
 FT /standard_name= "Single nucleotide polymorphism"
 FT replace (1097,G)
 FT /tag= d
 FT /standard_name= "Single nucleotide polymorphism"
 FT replace (1620,A)
 FT /tag= e
 FT /standard_name= "Single nucleotide polymorphism"
 FT replace (2119,C)
 FT /tag= f
 FT /standard_name= "Single nucleotide polymorphism"
 FT W0200231163-A1.
 FT 18-APR-2002.
 FT 11-OCT-2001; 2001MO-JP08913.
 FT 11-OCT-2000; 2000JP-0311309.
 FT 02-APR-2001; 2001JP-0102905.
 FT (KAZU-) KAZUSA DNA RES INST FOUND.
 FT (MITS-) MITSUBISHI PHARMA CORP.
 FT Ohara O, Nagase T, Nomura N, Yano K, Murakami K, Yasuda S;
 FT Kanzaki K;
 FT WPI; 2002-372227/40.
 FT P-PSDB; AAU79497.
 FT Human brain-originated ADAMTS family polypeptide and encoded gene,
 FT applicable in diagnosis and screening compounds for drug compositions
 FT in treating diseases due to e.g. neovascularisation
 FT Claim 4; Page 135-145; 172pp; Japanese.
 FT The invention relates to a polypeptide belonging to the ADAMTS family is
 FT selected from sequences appearing as AAU79496, AAU79497 and AAU79499,
 FT a protein that contains the polypeptide, a protein having not less than
 FT 5% homology with the amino acid sequence of the polypeptides or a
 FT polypeptide modified from any of the polypeptides but with some amino
 FT acids deleted, substituted, added or inserted. Also included are the
 FT polynucleotides encoding the polypeptides or their complementary
 FT strands or antisense, a vector construct containing any of the
 FT polynucleotides, a pharmaceutical which is administered to a
 FT patient, a process for producing the polypeptide, protein or peptide by culturing the
 FT transformant, an antibody that can recognize the polypeptide, protein or
 FT peptide and screening compounds to promote or inhibit activity of the
 FT polypeptide or protein, or to promote or inhibit expression of the
 FT polynucleotide by using the polypeptide, protein, peptide, the
 FT polynucleotide, vector, transformant or/and antibody, particularly in
 FT activity by measuring signal changes. The polypeptide and encoded gene
 FT are applicable in diagnosis and screening compounds for drug compositions
 FT in treating diseases due to neovascularisation, diabetic omentopathy,
 FT chronic rheumatoid arthritis, angiona, refractory skin and gastric ulcers
 FT and post-operative healing failure, including gene therapy.
 FT The gene encoding such polypeptide has conserved repressin-type 2N-
 FT type 1) domain. Its encoded protein is characterised by high expression
 FT in ovaries, changes in expression dose depending on the sexual cycle, a
 FT decrease in tumour cell and location of the gene on the 5p-syndrome
 FT Cri-du-chat syndrome). The present sequence encodes a ADAMTS
 FT protein of the invention.

Sequence 5610 BP; 1233 A; 1638 C; 1537 G; 1202 T; 0 other;

Query Match 64.9%; Score 1877.8; DB 24; Length 5610;
 Best Local Similarity 84.4%; Pred. No. 0;
 Matches 2289; Conservative

QY 202 GCAGCGACCTCGTGGCGCATGGAGCCGCGAGCGCGCTGGGAGCCCGAGCGTCCC 261
 DB 838 GCAGCGACCTCGTGGCGCATGGAGCCGCGAGCGCGCTGGGAGCCCGAGCGTCCC 897
 QY 262 GCGTCTCTCTCCACCCGCGAGCGCGCGGCTGGATGAGAAAGCGCGGACATGGATGA 321
 DB 898 GCGTCTCTCTCCACCCGCGAGCGCGCGGCTGGATGAGAAAGCGCGG----- 944
 QY 322 GCTGAAACCATCTCTTCTCAGCAAACTAACACAGGAACAGAAAACCAAAACATGCATGT 381
 DB 945 ----- 944
 QY 382 CTCACATATGACCTGTCTCTGCTACGAGGTTGACACAGGCGCGATTAGCTGTCTC 441
 DB 945 -----AATATGACCTGTCTCTGCTACGAGGTTGACACAGGCGCGATTAGCTGTCTC 997
 QY 442 CATGAATCATGCACCATCAGCGCGGAGAGAGAGAGTGGCGTGTCCGAGCTTGAGTCT 501
 DB 998 CATGAATCATGCACCATCAGCGCGGAGAGAGAGTGGCGTGTCCGAGCTTGAGTCT 1057
 QY 502 CTTCACTTCGGCTGGAAGGCGCCGAGGACGACTTCCATGAGATCTGAGGACTTCCAGC 561
 DB 1058 CTTCACTTCGGCTGGAAGGCGCCGAGGACGACTTCCATGAGATCTGAGGACTTCCAGC 1117
 QY 562 AGCTAGTGGCTCTCTGGCTTTATTTGTGCAGAGCTTGGGAAAGACAGGCACTAAGTCTGTG 621
 DB 1118 AGCTAGTGGCTCTCTGGCTTTATTTGTGCAGAGCTTGGGAAAGACAGGCACTAAGTCTGTG 1177
 QY 622 CAGACTTTACCGCCAGAGGACTTCTGTTTCTATCAAGGCTCTTTGGCATCTACACAGAAAC 681
 DB 1178 CAGACTTTACCGCCAGAGGACTTCTGTTTCTATCAAGGCTCTTTGGCATCTACACAGAAAC 1237
 QY 682 TGGCATTCGATCGAGGAGGAGTCTGTGAGGCGTCCACTGGCACTCTGAGCTTGGCAAC 741
 DB 1238 TCTTCA-----GTGGCTCTTTCACCTGCCAGGCTGTCTCA 1273
 QY 742 AGTCAGAAATGTCCCGGAGACAGTGTGACTTCTGCTGTCTGCTGCTGCGAGCAGACAC 801
 DB 1274 GGCATGATAGAACAGAGAGGAGATTAATCTTAAGGCCACTTCTTCAACACTCTCA 1333
 QY 802 AGCAGAGGATTGAGAGGCGGCGACTACAGTGG--AGCTTTACACTCAAGTAGAAGCG 859
 DB 1334 TGGAACTCGGAGAGTCCCGAGGAGCTCGCCATCCAGCTGACTGTACAGAGATCC 1393
 QY 860 ACTTATGCAACTTACTGTATCGCAGAGAGTGAATGTTCTTCTTTCTTTTCTTCAATA 919
 DB 1394 ALCAGGCCCATCTCTCTGGGCGCAGTGGTCTGTGACCTC-----AAGAT 1441
 QY 920 AGCTCAAGAGGAGTCTGCTGCTGCGGAGAGTACGCTGTTTGTATAGTAGTAGT 979
 DB 1442 ACATGGGCGCTGGCAGCATCAACCCCTGCACAGCAGGAGACTTCCGCTGGGAGTCCACAA 1501
 QY 980 GTGAGCTGTGTGTGTCCATCTGCG--GCACATGCTCCGAGCTCCCAAGGAAGACTC 1038
 DB 1502 AAGCAGCATTTCTGTGAGAGCGCAAGAAATACATGCTCCGAGCTCCCAAGGAAGACTC 1561
 QY 1039 TTCATCTCCGAGATGAGTATAGTCTTCTGCTAGGCATTAAGCGCTCTCTTCTGAGGTCC 1098
 DB 1562 TTCATCTCCGAGATGAGTATAGTCTTCTGCTAGGCATTAAGCGCTCTCTTCTGAGGTCC 1621
 QY 1099 CATGAATGAGAACTGAAGCTGGAGAGCTTGTGGTGTGTCGACAAAAGATGACAA 1158
 DB 1622 CATGAATGAGAACTGAAGCTGGAGAGCTTGTGGTGTGTCGACAAAAGATGACAA 1681
 QY 1159 AACCATGGCCATGAATAATCAACCACTACGTGCTCAACGATCTCAACATGATCTGCT 1218
 DB 1682 AACCATGGCCATGAATAATCAACCACTACGTGCTCAACGATCTCAACATGATCTGCT 1741
 QY 1219 TTATTCAAAGA----- 1229

DB	2822	CGAAGAGGATTTGGATCTCTTTCTTGTGCAATTAAGAGTCAAAGCATGGGCACTCCCATGC	2883
QY	2233	TCGGAGGATAGCGTAAATGTTCTATAGCATGGGTAATGVSAGAGGCTGGGATGTGACAAAT	2292
DB	2882	TCGGAGGATAGCGTAAATGTTCTATAGCATGGGTAATGVSAGAGGCTGGGATGTGACAAAT	2941
QY	2293	GTCTCTTGAGATCTGATGCTGTGTGAAGAGGCTCTGTGGGGGTGTCTAACGGGNAATAACTCAGCC	2352
DB	2942	GTCTCTTGAGATCTGATGCTGTGTGAAGAGGCTCTGTGGGGGTGTCTAACGGGNAATAACTCAGCC	3001
QY	2353	TGCACGATTTCCAGAGGGGTCTCTACACCAAGCACACCAACACCACTGGACCTTATCATGATG	2412
DB	3002	TGCACGATTTCCAGAGGGGTCTCTACACCAAGCACACCAACACCACTGGACCTTATCATGATG	3061
QY	2413	GTACACCAATTCCTCTGSGAGCCCGGAGTATCCGCATCTATGAAATGAAAGCTCTCTACCTCC	2472
DB	3062	GTACACCAATTCCTCTGSGAGCCCGGAGTATCCGCATCTATGAAATGAAAGCTCTCTACCTCC	3121
QY	2473	TACATTTCTGTGGCGCAATGCCCTCAGAAGGTACTACTGAATGGGCACCTGGACCCGTGGAC	2532
DB	3122	TACATTTCTGTGGCGCAATGCCCTCAGAAGGTACTACTGAATGGGCACCTGGACCCGTGGAC	3181
QY	2533	TGGCCCGCGCCGTACAAATTTTTCGGGCACTACTTTCGACTACAGAGGTCTCTAATATGAG	2592
DB	3182	TGGCCCGCGCCGTACAAATTTTTCGGGCACTACTTTCGACTACAGAGGTCTCTAATATGAG	3241
QY	2593	CCCGAGAACTTAATTCGCTACTTGGACCAACCAAGACACACTGATTTGGAGCTGCTGTTT	2652
DB	3242	CCCGAGAACTTAATTCGCTACTTGGACCAACCAAGACACACTGATTTGGAGCTGCTGTTT	3301
QY	2653	CAGGGAAGAAACCGGGGTGTGTGCTTGGGATACTTCCATGCTCTGGGACCGGAGAG	2712
DB	3302	CAGGGAAGAAACCGGGGTGTGTGCTTGGGATACTTCCATGCTCTGGGACCGGAGAG	3361
QY	2713	CAGCCCCCTGTGCCAGGCCAGCTACACTTGGGCGCATGCTGGCGCTCTGAGTGGCTCGGTGCC	2772
DB	3362	CAGCCCCCTGTGCCAGGCCAGCTACACTTGGGCGCATGCTGGCGCTCTGAGTGGCTCGGTGCC	3421
QY	2773	TGCGGAGGGGG 2783	
DB	3422	TGCGGAGGGGG 3432	
DB	RESULT 5		
DB	ABK49825		
ID	ABK49825	standard; DNA, 8435 BP.	
XX	AC		
XX	ABK49825;		
DT	15-JUL-2002	(first entry)	
DE	Plasmid pFascBac1-HF-P001256.		
XX	ds, ADAMTS; cytostatic; antidiabetic; antirheumatic; cyclic;		
XX	antiarthritic; anticulcer; vulnery; neovascularisation; angiona;		
XX	diabetic omentopathy; chronic rheumatoid arthritis; gene therapy;		
XX	refractory skin ulcer; gastric ulcer; post-operative healing failure;		
XX	reprolysin-type Zn-metalloprotease domain; disintegrin-like domain; TSPI;		
XX	chromosomid type domain; sexual cycle; tumour; 5p-syndrome deletion;		
XX	chromosome SP3; 15.3; Cri-du-chat syndrome; circular;		
XX	pFascBac1-HF-P001256.		
XX	OS Homo sapiens.		
XX	OS Escherichia coli.		
XX	OS Baculovirus.		
XX	OS Rhesus macaque polyoma virus.		
XX	OS Tussock moth.		
XX	OS Synthetic.		
XX	WC200231163-A1.		
PN	18-APR-2002.		
PD			

Db	4280	CATGAAATCATGCAACCATCAGCGCGGAGAGAGACGATGGCGGTGTCGAGGTTGAAGTCT	4331
Qy	502	CTTCACCTTTGGCTGAAAGCGCCCGAGCAGACATTCACATGGATCTGAGGACTTCACGC	561
Db	4340	CTTCACCTTTGGCTGAAAGCGCCCGAGCAGACATTCACATGGATCTGAGGACTTCACGC	4399
Qy	562	AGCCTAGTGGCTCTCGCTTTATGTGCGACAGCTTGGGAAGAAGCAGGCACTAAGTCTGTG	621
Db	4400	AGCCTAGTGGCTCTCGCTTTATGTGCGACAGCTTGGGAAGAAGCAGGCACTAAGTCTGTG	4459
Qy	622	CAGACTTTACCGCCAGAGGACCTCTGTTTCTATCAAGGCTCTTTTGGGATCACAAGAAAC	681
Db	4460	CAGACTTTACCGCCAGAGGACCTCTGTTTCTATCAAGGCTCTTTTGGGATCACAAGAAAC	4519
Qy	682	TGSCCATTCGCATGAGGGAAGTTCTGTGAGGGCTCCACTGCGACCTCTGAAAGCTCTCGAC	741
Db	4520	TGCTCA-----GTGGCCCTTTCAAGCTGCCAAGGCTTGTC	4555
Qy	742	AGTCAGAAATGTCCCGGGCAGAGTGTGACTTCCTGGTGTCTGCTGCTCGAGTGTGCGAGCACAAC	801
Db	4556	GGCATGATACGAAAGAAAGGCGAGATTACTTCTTAAGGCCACTTCCTTCACACCTCTCA	4615
Qy	802	AGCAGACGATTCAGAGGGCGGCACCTACAAGTGG--AAGCCTTTACATCTCAAGTAGAGCGG	859
Db	4616	TGGAATCTGGCAGAGCTCCCAAGGCACTGGCCATCCCAAGTCTGTACAAGAGATCC	4675
Qy	860	ACTTATGCAAACTTACTGTATCTCGAGAAGGATTGATTTCTTTCTTTCTTTTGTCAATA	919
Db	4676	ACAGAGCCCCCATGCTCTCTGGGGCAGTGAGGTCTCTGTGAAGCTC-----AAGG	4723
Qy	920	AAGTCAAAATGGGACCTCCATGCTTCGAGAGATAGCGTAATGTGTTGTATAGATGGGATAT	979
Db	4724	ACATGGGAGCTGGCACAACAACCCCTCGACAGAGGACCTTCGCTCGGAGACTGCCACA	4783
Qy	980	GTGAGCTCGAGTGGTCCACATCTGC--GCACATGCCCCAGCCTCCCAAGGAAGACCTC	1038
Db	4784	AAGACGCAATTTCTGTGGAAGACGAGAATAATACATGCCCCAGCCTCCCAAGGAAGACCTC	4843
Qy	1039	TTCAATCTCGAGTAGATTAAGCTCTTCTTAGGGCAATAGCGCTCTCTTCTAGAGTCC	1098
Db	4844	TTCAATCTCGAGTAGATTAAGCTCTTCTTAGGGCAATAGCGCTCTCTTCTAGAGTCC	4903
Qy	1099	CAATCAATTAAGAACTGAAGCTGGAGACCTGGTGGTGGTCGCAAAAAGATGATGCA	1158
Db	4904	CATCAATTAAGAACTGAAGCTGGAGACCTGGTGGTGGTCGCAAAAAGATGATGCA	4963
Qy	1159	AACTGGCGCAAGAAATCAACATCACTGCTCTCAAGATCTCAACATGTGATCTCT	1218
Db	4964	AACTGGCGCAAGAAATCAACATCACTGCTCTCAAGATCTCAACATGTGATCTCT	5023
Qy	1219	TTATTCAAAG-----	1239
Db	5024	TTATTCAAAGATGGACATAGGAGGAACATCAACATTTGCATTTGTAGGTCTGATCTCT	5081
Qy	1230	-----	1239
Db	5084	CTAGAGATGACAGCCAGCATCGTGATTAAGTCAACCGCAGACACACCTTAAGTAGC	5143
Qy	1230	-----TGGATTCACTGGGGAAGATGGGACTCTGCTATGACCAACCGCCCTC	1272
Db	5144	TTCTGCCAGTGGCAGTCTGGATTGCTGGGGAAGATGGGACTCTGCTATGACCAAGCCATC	5203
Qy	1273	TTACTGACTGGTCTGATATATGTTCTGGAAGAATAGCCCTCTGAGACATTTTGGGATTT	1332
Db	5204	TTACTGACTGGTCTGATATATGTTCTGGAAGAATAGCCCTCTGAGACATTTTGGGATTT	5263
Qy	1333	GCACCCATAAGTGGAAATGTGTAGTAATAATGCGAGCTGCAAGATTAATGAGATACAGGT	1392
Db	5264	GCACCCATAAGTGGAAATGTGTAGTAATAATGCGAGCTGCAAGATTAATGAGATACAGGT	5323
Qy	1393	CTTGACTGGCCCTTACCAATGGCCCATGTAGTCTGGACACAACTTTGGCAAGTCAATGAT	1452
Db	5324	CTTGACTGGCCCTTACCAATGGCCCATGTAGTCTGGACACAACTTTGGCAAGTCAATGAT	5383

QY 1453 GCGAGGCGACATGTGTAAAGAGTCGAGGCGCAACATCATGTCCCTACATATGCGACGA 1512
 DB 5384 GCGAGGCGACATGTGTAAAGAGTCGAGGCGCAACATCATGTCCCTACATATGCGACGA 5443
 QY 1513 CGCAATGAGAGTCTCTCTGCTGATCAGGCGGCGGAGTATCTACACAAATTTCTAGC 1572
 DB 5444 CGCAATGAGAGTCTCTCTGCTGATCAGGCGGCGGAGTATCTACACAAATTTCTAGC 5503
 QY 1573 ACCTCTCAAGCTATCTGCTGCTGATCAGGCGGCGGAGTATCTACACAAATTTCTAGC 1632
 DB 5504 ACCTCTCAAGCTATCTGCTGCTGATCAGGCGGCGGAGTATCTACACAAATTTCTAGC 5563
 QY 1633 GAGAAATTTGCCAGAGAGATTTATATGTCAGAACACACAGTGTGAGTGTGCGAGAG 1692
 DB 5564 GAGAAATTTGCCAGAGAGATTTATATGTCAGAACACACAGTGTGAGTGTGCGAGAG 5623
 QY 1693 AAAGCCAGAGCTCTGCTGCTGATCAGGCGGCGGAGTATCTACACAAATTTCTAGC 1752
 DB 5624 AAAGCCAGAGCTCTGCTGCTGATCAGGCGGCGGAGTATCTACACAAATTTCTAGC 5683
 QY 1753 GGTATTTGAGAGAGATTTGAGAGTAAATTTATGCGAGAGAGAGAGAGAGAGAGAGAG 1812
 DB 5684 GGTATTTGAGAGAGATTTGAGAGTAAATTTATGCGAGAGAGAGAGAGAGAGAGAGAG 5743
 QY 1813 CATGACATGTGTCGCGGAGAGAGAGAGTGTGTGAATATGTGTATGAGAGGCGGCGGAG 1872
 DB 5744 CATGACATGTGTCGCGGAGAGAGAGTGTGTGAATATGTGTATGAGAGGCGGCGGAG 5803
 QY 1873 ACCATGCGGAGTGTGTCGAGTGTGTGTCCTGTCGAGAGAGAGAGAGAGAGAGAGAG 1932
 DB 5804 ACCATGCGGAGTGTGTCGAGTGTGTGTCCTGTCGAGAGAGAGAGAGAGAGAGAGAG 5863
 QY 1933 GAGATTTCTATGAGAGTGTGTCGAGTGTGTGTCCTGTCGAGAGAGAGAGAGAGAG 1992
 DB 5864 GAGATTTCTATGAGAGTGTGTCGAGTGTGTGTCCTGTCGAGAGAGAGAGAGAGAG 5923
 QY 1993 GTGAGAGGCTCTGAGTGTGTCGAGTGTGTGTCCTGTCGAGAGAGAGAGAGAGAG 2052
 DB 5924 GTGAGAGGCTCTGAGTGTGTCGAGTGTGTGTCCTGTCGAGAGAGAGAGAGAGAG 5983
 QY 2053 GTTGAATTTCTGTCGTCAGTGTGTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2112
 DB 5984 GTTGAATTTCTGTCGTCAGTGTGTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6043
 QY 2113 TACAAGTGAAGGCTTACACTCAAGTAGAAGATCAGAGCTTATGCAAACTCTACTGTATC 2172
 DB 6044 TACAAGTGAAGGCTTACACTCAAGTAGAAGATCAGAGCTTATGCAAACTCTACTGTATC 6103
 QY 2173 GCGAGAGGATTTGATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2232
 DB 6104 GCGAGAGGATTTGATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 6163
 QY 2233 TCGAGAGATGAGCCGATTAATGTTTGTATAGATGGATATGTGAGAGAGTGTGATGTGACAT 2292
 DB 6164 TCGAGAGATGAGCCGATTAATGTTTGTATAGATGGATATGTGAGAGAGTGTGATGTGACAT 6223
 QY 2293 GTCTTGTGATGATGCTGTTGAGAGAGTCTGTGGGTTGTGTAACCGGGAATACTCAGCC 2352
 DB 6224 GTCTTGTGATGATGCTGTTGAGAGAGTCTGTGGGTTGTGTAACCGGGAATACTCAGCC 6283
 QY 2353 TGCAAGATTCAGAGGCTCTCTACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2412
 DB 6284 TGCAAGATTCAGAGGCTCTCTACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6343
 QY 2413 GTCAAGATTCCTTCTGAGGCGGAGTATGCGATCTATGAAATGAACTCTCTACCTCC 2472
 DB 6344 GTCAAGATTCCTTCTGAGGCGGAGTATGCGATCTATGAAATGAACTCTCTACCTCC 6403
 QY 2473 TACATTTCTGCGGCAATGCTCTCAGAGAGTACTCTCTGATGAGGAGTGTGAGAGGAGAG 2532
 DB 6404 TACATTTCTGCGGCAATGCTCTCAGAGAGTACTCTCTGATGAGGAGTGTGAGAGGAGAG 6463

QY 2533 TGCGCGGCGCGGTACAAATTTTCGGGCACTACTTTTCGACTACAGACGGTCTCTATATGAG 2592
 DB 6464 TGCGCGGCGCGGTACAAATTTTCGGGCACTACTTTTCGACTACAGACGGTCTCTATATGAG 6523
 QY 2593 CCGGAGAGACTTAATCCCTACTGAGACCCACCAAGAGACACTGATTTGAGAGCTGCTGTTT 2652
 DB 6524 CCGGAGAGACTTAATCCCTACTGAGACCCACCAAGAGACACTGATTTGAGAGCTGCTGTTT 6583
 QY 2653 CAGGAGAGAAACCGGGGTGTTGCTTGGGATACCTTCCATGCTCTGCTTGGGAGACCGAGAG 2712
 DB 6584 CAGGAGAGAAACCGGGGTGTTGCTTGGGATACCTTCCATGCTCTGCTTGGGAGACCGAGAG 6643
 QY 2713 CAGGCGGCTTGGCCAGACCCAGCTACACTTGGGCGCATCTGTCGCTCTGAGTGTCTGCTGTC 2772
 DB 6644 CAGGCGGCTTGGCCAGACCCAGCTACACTTGGGCGCATCTGTCGCTCTGAGTGTCTGCTGTC 6703
 QY 2773 TGCGGAGGGGG 2783
 DB 6704 TGCGGAGGGGG 6714

RESULT 6
 ABK49826
 ID ABK49826 standard; DNA; 8505 BP.
 XX AC ABK49826;
 XX DT 15-JUL-2002 (first entry)
 XX DE Plasmid pFastBac1-MS/HT-F01256-2.
 XX ds; ADAMTS, cytostatic; antidiabetic; antirheumatic; cyclic;
 XX antiarthritic; antitumor; antineoplastic; antineoplastic; antineoplastic; antineoplastic;
 XX diabetic omentopathy; chronic rheumatoid arthritis; gene therapy;
 XX retractor skin (cat); gastric ulcer; post-operative healing failure;
 XX psoriasis; type 1 diabetes; sexual cycle; tumor; 5p-syndrome deletion;
 XX chromosome 2-15.3 Cri-du-chat syndrome; circular;
 XX pFastBac1-MS/HT-F01256-2.
 XX Homo sapiens.
 OS Escherichia coli.
 OS Baculovirus.
 OS Rhesus macaque polyoma virus.
 OS Transposon Tn7.
 OS Synthetic.
 XX WO200231163-A1.
 XX 18-APR-2002.
 XX 11-OCT-2001; 2001WO-JP08913.
 XX 11-OCT-2000; 2000JP-0311309.
 XX 02-APR-2001; 2001JP-0102905.
 XX (KAZU-) KAZUSA DNA RES INST FOUND.
 XX (MITS-) MITSUBISHI PHARMA CORP.
 XX Ohara O, Nagase T, Nomura N, Yano K, Murakami K, Yasuda S;
 XX Kanzaki K;
 XX WPI; 2002-372277/40.
 XX Human brain-originated ADAMTS family polypeptide and encoded gene,
 XX applicable in diagnosis and screening compounds for drug compositions
 XX in treating diseases due to e.g. neovascularisation.
 XX Example 20; Page 159-163; 172pp; Japanese.
 XX The invention relates to a polypeptide belonging to the ADAMTS family is
 XX selected from sequences appearing as AAU79496, AAU79497 and AAU79499,
 XX a protein that contains the polypeptide, a protein having not less than

DB 289 CTTCACTTGGCTGMAAGGCTCCAGGACGACTTCCACGTGATCTGAGGACTTCCAGC 348
 QY 562 AGCTATGTCGCTGCTGCTTTATTTGACAGCTTGGAAAGACAGGACCTAAGTCTGTG 621
 DB 349 AGCTATGTCGCTGCTGCTTTATTTGACAGCTTGGAAAGACAGGACCTAAGTCTGTG 408
 QY 622 CAGACTTTACGACAGGACCTGCTGCTTTATTTGACAGCTTGGAAAGACAGGACCTA 681
 DB 409 CAGCTTTACGACAGGACCTGCTGCTTTATTTGACAGCTTGGAAAGACAGGACCTA 468
 QY 682 TCGGCTATCGATGAGGAGGACTTGTGAGGCTCCACTCGCACTCTGAGCTTGCAC 741
 DB 469 TCTCTCA-----GTGGCCCTTCAAGCTCGAGGCTTGCA 504
 QY 742 AGTCAGAACTCCCGGACAGTGTCACTTCGCTGCTGCTCACTGTGCGAGCAGAAC 801
 DB 505 GGCATGATACGACAGAGAGGACGATTACTTCTTAAGGCCACTTCTTCAACACCTTCA 564
 QY 802 AGCAGAGATTCAGAGGGGCGCACTACAGTGG--AAGCTTTACACTCAGTACAGAGCCG 859
 DB 565 TGGAACTTCGACAGCTGCCAAGGAGCTGCCCATCCACTGTACAGAGATCC 624
 QY 860 ACTTATGCAAACTTACTGATGCGAGAGGATTTGATTTCTTTCTTTCTTTCTCAATA 919
 DB 625 ACAGAGCCCATGCTCTGCGGCGCAGTGAGTCTTGGTGCCTC-----AAGG 672
 QY 920 AAGTCAAGATGAGGACTCCATGCTCGAGGATAGCCGTAATGTTGTATAGATGGGATAT 979
 DB 673 ACATGGGAGCTGCGACATCAACCCCTGACAGAGCGACCTTCGCTCGGACTGCGACAA 732
 QY 980 GTGAGCTCAGTGTGTGTCACATCTGC--GCACATGCCCGAGCCTCCCAAGGAAGCTC 1038
 DB 733 AAGCAGCATTTCTGTGGAGGCGCAGAAATATCATGCGCCAGCTCCCAAGGAGACCTC 792
 QY 1039 TTCACTTTCGACAGATGATTAAGTCTTCTTACGCGATTAAGGCTCTCTCTTCTAGTTC 1098
 DB 793 TTCACTTTCGACAGATGATTAAGTCTTCTTACGCGATTAAGGCTCTCTCTTCTAGTTC 852
 QY 1099 CHTAGAAATGAGAACTGAACTGGGAGCTTGGTGTGCTGCTGCTGCTGCTGCTGCTGCT 1158
 DB 853 CHTAGAAATGAGAACTGAACTGGGAGCTTGGTGTGCTGCTGCTGCTGCTGCTGCTGCT 912
 QY 1159 AACGTCGCTGATGAAATATCAACCTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1218
 DB 913 AACCTTGGCCATGAAATATCAACCTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 972
 QY 1219 TTATTCNAGR----- 1229
 DB 973 TTATTCNAGATGGAACATAGGAGGAAACATCAACATTTGCAATTTGATGCTGATCTT 1032
 QY 1230 ----- 1229
 DB 1033 CTAGAAGATGAACCGAGGACTGCTGATTAAGTCAACACGACACACACTTTAAGTAGC 1092
 QY 1236 -----TGGAATGATGGGAAAGATGAGGACTGCTCATGACACCGCAATC 1272
 DB 1093 TTCTGCGAGTGGCAGTCTGGAATTTGATGGGAAAGATGAGGACTGCTCATGACACCGCAATC 1152
 QY 1273 TTATGATGCTGCTGGAATATGTTCTGGAAGAAATGAGCCTGTGACACTTTGGGATTT 1332
 DB 1153 TTATGATGCTGCTGGAATATGTTCTGGAAGAAATGAGCCTGTGACACTTTGGGATTT 1212
 QY 1333 GCACCAATAGTGGATGTAGTAATATCGAGCTGACAGCTGACAGATTAATGAGATACAGGT 1392
 DB 1213 GCACCATTAAGTGGATGTAGTAATATCGAGCTGACAGCTGACAGATTAATGAGATACAGGT 1272
 QY 1393 CTTGGAATGCTGCTGCAATGTCGAGCTGGAACAACTTTGGCATGATTAATGAT 1452
 DB 1273 CTTGGAATGCTGCTGCAATGTCGAGCTGGAACAACTTTGGCATGATTAATGAT 1332
 QY 1453 GGAAGAGGACATGTTAAAGTCCAGGGGCACTATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1512
 DB 1333 GGAAGAGGACATGTTAAAGTCCAGGGGCACTATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1392

PT Preparation of a medicament for diagnosing, preventing or treating
 disorders or diseases associated with the NHP, e.g., obesity
 XX Disclosure; Column 25-30; 17pp; English.

XX The invention describes a new isolated nucleic acid comprising a
 sequence that encodes a fully defined protein sequence comprising 491
 or 1224 amino acids or that hybridises under stringent conditions with
 the 3675-bp sequence of its complement. The nucleic acid is useful for
 CC the preparation of a medicament for diagnosing, preventing or treating
 CC disorders (e.g. using gene therapy) or diseases associated with the
 CC novel human proteins, e.g., obesity. This sequence encodes a novel
 CC human protease.

XX Sequence 4042 BP; 947 A; 1117 C; 1130 G; 848 T; 0 other;

Query Match 64.8%; Score 1874.6; DB 25; Length 4042;
 Best Local Similarity 84.4%; Pred. No. 0;
 Matches 2287; Conservative 0; Mismatches 179; Indels 245; Gaps 6;
 QY 202 GAGGACCTGTGGCCGATGAGACCGGACGCGAGCGCTTGGAGCCCGAGGCTCC 261
 DB 167 GCGGACCTGTGGCCGATGAGACCGGACGCGAGCGCTTGGAGCCCGAGGCTCC 226
 QY 262 GGTCTCTCTCCACCCGCGGAGCGCGCGCTGGATGGAAAGCGCGGACATGGTGA 321
 DB 227 GGTCTCTCTCCACCCGCGGAGCGCGCGCTGGATGGAAAGCGCGGACATGGTGA 273
 QY 322 GCTGGAAACATGCTTCTGACAACTAACACAGGACAGAAAACCAACATGCTGTT 381
 DB 274 ----- 273
 QY 382 CTCACATGATGAGCTGTCTCTGCTACGAGTTGACACAGAGCGGATTAAGTGTCC 441
 DB 274 -----NATNTGACCTGTCTCTCTGCTACGAGTTGACACAGAGCGGATTAAGTGTCC 326
 QY 442 CATGAATCATGACCATCAGCGCGAGGAGAGCAGTGGCGTCCGAGGTGAGTCT 501
 DB 327 CATGAATCATGACCATCAGCGCGAGGAGAGCAGTGGCGTCCGAGGTGAGTCT 386
 QY 502 CTTGACCTTGGCTGAAAGCCCGACAGGACGATCTCCACATGATCTGAGGACTTCCAG 561
 DB 387 CTTGACCTTGGCTGAAAGCCCGACAGGACGATCTCCACATGATCTGAGGACTTCCAG 446
 QY 562 AGCTTAGTGGCTCTGGCTTTATTTGTGAGAGCGTTGGGAAAGACAGGACCTAAGTCTGT 621
 DB 447 AGCTTAGTGGCTCTGGCTTTATTTGTGAGAGCGTTGGGAAAGACAGGACCTAAGTCTGT 506
 QY 622 CAGACTTTACCGCAGAGGACTTGTGTTCTATCAAGGCTCTTTCGATCACAGAAAC 681
 DB 507 CAGACTTTACCGCAGAGGACTTGTGTTCTATCAAGGCTCTTTCGATCACAGAAAC 566
 QY 682 TCGCATCGCATGGAGGAAGTCTGTGGGGCTCCACTCGACTCTGAAGCTCTGCAAC 741
 DB 567 TCGTCA-----GTGGCCCTTTCAACCTGCCAAGGCTTGTCA 602
 QY 742 AGTCAGAAATGTCCTCCCGGACAGTGTGACTTCTGCTGCTGCTGAGTGGCCGACACAG 801
 DB 603 GGCATGATGACAGAGAGAGGACAGATTAATCTCTAAGGGCACTTCTCTCACACCTTCA 662
 QY 802 AGCAGAGATTCAGAGGCGGCGCTACAGTGG--AAGCTTTACACTCAAGTAGAGCGG 859
 DB 663 TGGAAACTCGGCGAGGCTGCGCCAGGCGAGCTGCGCATCCCACTACTGTACAGAGATCC 722
 QY 860 ACTTATGCAACTCTACTGTATCGGAGAGGATTTGATTTCTTTCTTTCTTTGTCAAATA 919
 DB 723 ACAGAGCCCATGCTCTCTGGGCGAGTGGAGTCTCTGTGAGCTC-----AAGG 770
 QY 920 AGTCAGAGGAGGAGTCTCTGCTGAGGAGTAGCGTAAATGTTGTATGATGAGATAT 979
 DB 771 ACATGGAGCTGAGACATCAACCTCTCAACAGCGACCTTCGCTGGAGCTGCAAA 830
 QY 980 GTGAGCTCAGTGTGTGTCACATCTGC-GCATGTGCCCGCCAGCTCCCAAGAGAACCTC 1038

DB 831 AAGAGCATTCTTGTGGAGACGCAAGAAATACATGCCCCAGCCTCCCAAGAGAGCCTC 890
 QY 1039 TTCTATCTGGCAGATGAGTATAGTCTTGTCTAGCGTAAAGCGCTCTCTCTAGGCTCC 1098
 DB 891 TTCTATCTGGCAGATGAGTATAGTCTTGTCTAGCGTAAAGCGCTCTCTCTAGGCTCC 950
 QY 1099 CATAGAATGAAGAACTGAACTGGAGACCTTGTGGTGTGTGCGACAAAGATGATGACAA 1158
 DB 951 CATAGAATGAAGAACTGAACTGGAGACCTTGTGGTGTGTGCGACAAAGATGATGACAA 1010
 QY 1159 AACCATGCCCATGAAATATATCAACCTACCTAGCTGTCTACAGTACTCAACATGATGCT 1218
 DB 1011 AACCATGCCCATGAAATATATCAACCTACCTAGCTGTCTACAGTACTCAACATGATGCT 1070
 QY 1219 TTATTTAAAGA----- 1229
 DB 1071 TTATTTAAAGATGGAAACATATAGGAGAAACATCAACATTTGATAGTCTGTATCTT 1130
 QY 1230 ----- 1229
 DB 1131 CTAGAAGATGACACCGACAGACCTGTGTATGTACACACGACACACCTTTAAGTAGC 1190
 QY 1230 -----TGGATGTGGGAAAGATGGAGCTGTGTATGTACACACGACACCTTTAAGTAGC 1272
 DB 1191 TTCTGCCAGTGGACGCTGTGTATGTATGGGAAAGATGGAGCTGTGTATGTACACACCT 1250
 QY 1273 TTTACTGACTGTGTGATATATGTCTGTGAAAGATGGAGCTGTGTACACTTTGGGATTT 1332
 DB 1251 TTTACTGACTGTGTGATATATGTCTGTGAAAGATGGAGCTGTGTACACTTTGGGATTT 1310
 QY 1333 GCACCATTAAGTGTGTGTGTATATTCGACGCTGTGACGCTTATTAAGTGTACAGT 1392
 DB 1311 GCACCATTAAGTGTGTGTGTATATTCGACGCTGTGACGCTTATTAAGTGTACAGT 1370
 QY 1393 CTTGAGCTGSCCTTCAACCTTGGCCATGAGTGTGACACACTTTGGCATGATTCATGAT 1452
 DB 1371 CTTGAGCTGSCCTTCAACCTTGGCCATGAGTGTGACACACTTTGGCATGATTCATGAT 1430
 QY 1453 GAGGAGGAGAACTGTGTAAGAGTTCGAGGCGACATCATGTCCCTTACATTTGGCAGGA 1512
 DB 1431 GAGGAGGAGAACTGTGTAAGAGTTCGAGGCGACATCATGTCCCTTACATTTGGCAGGA 1490
 QY 1513 CGCATGAGTCTTCTCTGTGCTGATCAGCCGACGCTGTATACACAAATTTCTAAGC 1572
 DB 1491 CGCATGAGTCTTCTCTGTGCTGATCAGCCGACGCTGTATACACAAATTTCTAAGC 1550
 QY 1573 ACCGCTCAAGCTTATCTGCTTGTGATCAGCCGACGCTGTGAGAGGATACAGTATCTCT 1632
 DB 1551 ACCGCTCAAGCTTATCTGCTTGTGATCAGCCGACGCTGTGAGAGGATACAGTATCTCT 1610
 QY 1633 GAGAAATTTGCGAGAGAAATATATGATGACAAACACAGTGTGAGTGGAGTTCGAGAG 1692
 DB 1611 GAGAAATTTGCGAGAGAAATATATGATGACAAACACAGTGTGAGTGGAGTTCGAGAG 1670
 QY 1693 AAAGCAAAGCTTCTGCTGCTGAGCTTTAAAGAGGACATCTGTAAGAGCCCTGTGTGCGCAT 1752
 DB 1671 AAAGCAAAGCTTCTGCTGCTGAGCTTTAAAGAGGACATCTGTAAGAGCCCTGTGTGCGCAT 1730
 QY 1753 CGTATTGGAAGGAAATGTGAGACTAAATTTATGCCAGCAGCAGAGAGGACAAATTTGTGG 1812
 DB 1731 CGTATTGGAAGGAAATGTGAGACTAAATTTATGCCAGCAGCAGAGAGGACAAATTTGTGG 1790
 QY 1813 CATGACATGTGTGTCGCGGAGGACAGTGTGTGAAATATGCTGATGAAAGGCCCCCAAGCCC 1872
 DB 1791 CATGACATGTGTGTCGCGGAGGACAGTGTGTGAAATATGCTGATGAAAGGCCCCCAAGCCC 1850
 QY 1873 ACCCATGCGCACTGTGTGTGCTGTGCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1932
 DB 1851 ACCCATGCGCACTGTGTGTGCTGTGCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1910
 QY 1933 GAGTATCTCTATGAGGTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1992

Db 1911 GGAGTATCTCATAGAGTTCGCTCTTGCACACACCCCAAGCCATTCGATCGAGGGAAGTTC 1970
 Qy 1993 TGTGAGGCTTCACCTCGCATCTCTGAAGCTCTGCACACAGTCAGAAATGTCCTCCGGACAGT 2052
 Db 1971 TGTGAGGCTTCACCTCGCATCTCTGAAGCTCTGCACACAGTCAGAAATGTCCTCCGGACAGT 2030
 Qy 2053 GTTGACTTCGCTGCTCTGATGTGCTCCGACACACACAGCAGAGGATTCAGAGGGCGGCAC 2112
 Db 2031 GTTGACTTCGCTGCTCTGATGTGCTCCGACACACACAGCAGAGGATTCAGAGGGCGGCAC 2090
 Qy 2113 TACAGTGGAGGCTTACACTCACTCAAGTAGAGATCAGGACTTATGCAAACTCTACTGTATC 2172
 Db 2091 TACAGTGGAGGCTTACACTCACTCAAGTAGAGATCAGGACTTATGCAAACTCTACTGTATC 2150
 Qy 2173 GGNAGAGATTGATCTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2232
 Db 2151 GGNAGAGATTGATCTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2210
 Qy 2233 TGGAGGATAGCCGTAATGTTTGTATAGATGGGATATGTGAGAGATTTGATGTGACAT 2292
 Db 2211 TGGAGGATAGCCGTAATGTTTGTATAGATGGGATATGTGAGAGATTTGATGTGACAT 2270
 Qy 2293 GTCTTGTGATCTGATCTGTGTGAGAGCTCTGTGGGGTGTGTAAAGGATTAATCACTG 2352
 Db 2271 GTCTTGTGATCTGATCTGTGTGAGAGCTCTGTGGGGTGTGTAAAGGATTAATCACTG 2330
 Qy 2353 TGCAGATTCACAGGGGTCTTACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2412
 Db 2331 TGCAGATTCACAGGGGTCTTACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2390
 Qy 2413 GTACGATCTCTCTGAGCGCGGAGTATCGGATCTATGAAATGAAAGCTCTCACTCC 2472
 Db 2391 GTACGATCTCTCTGAGCGCGGAGTATCGGATCTATGAAATGAAAGCTCTCACTCC 2450
 Qy 2473 TACATTTCTGTGGCGCAATGCGCTCGAGAGGTACTACTGATGGGACCTGACCGGAC 2532
 Db 2451 TACATTTCTGTGGCGCAATGCGCTCGAGAGGTACTACTGATGGGACCTGACCGGAC 2510
 Qy 2533 TGGCGCGCGGTACAAATTTGGCGGCTACTTTGCACTAGAGAGGCTCTATTAACAG 2592
 Db 2511 TGGCGCGCGGTACAAATTTGGCGGCTACTTTGCACTAGAGAGGCTCTATTAACAG 2570
 Qy 2593 CCGGAGACCTTATCGCTACTGACACACAGACAGCTGATTTGGAGCTCTGTTT 2652
 Db 2571 CCGGAGACCTTATCGCTACTGACACACAGACAGCTGATTTGGAGCTCTGTTT 2630
 Qy 2653 CAGGGAAGGAACCGGCTGTGCTGCGGATATCTCGCTCGCTTGGGACCGAGAG 2712
 Db 2631 CAGGGAAGGAACCGGCTGTGCTGCGGATATCTCGCTCGCTTGGGACCGAGAG 2690
 Qy 2713 CAGCCCCCTGCCCGCCCACTACTCTGAGCCATCGTGGCTCTGAGTGTCTCGTGTCC 2772
 Db 2691 CAGCCCCCTGCCCGCCCACTACTCTGAGCCATCGTGGCTCTGAGTGTCTCGTGTCC 2750
 Qy 2773 TCGGAGGGGG 2783
 Db 2751 TCGGAGGGGG 2761

RESULT 10
ABSS9324

ID ABSS9324 standard; DNA; 2432 BP.

AC ABSS9324;

XX

DT 05-NOV-2002 (first entry)

DE Human zinc metalloprotease-like gene #2.

XX Human; NOVX; cardiomyopathy; atherosclerosis; cell signal processing;

KW breast cancer; Alzheimer's disease; epilepsy; Huntington's disease;

KW anxiety; behavioural disorder; multiple sclerosis; myasthenia gravis;

KW neurodegeneration; Parkinson's disease; pain; stroke; endometriosis;

KW autoimmune disease, allergy; addiction; asthma; transplantation;
 KW graft versus host disease; systemic lupus erythematosus; scleroderma;
 KW psoriasis; Crohn's disease; HIV infection; human immunodeficiency virus;
 KW atherosclerosis; cirrhosis; rheumatoid arthritis; diabetes; pancreatitis;
 KW thrombocytopenia; bleeding disorder; metabolic disorder; obesity;
 KW glucose transport defect; glomerulonephritis; hypercalcaemia;
 KW polycystic kidney disease; renal tubular acidosis; skin disorder;
 KW congenital diarrhoea; respiratory disease; gastro-intestinal disease;
 KW muscle disorder; bone disorder; joint disorder; skeletal disorder;
 KW haematopoietic disorder; urinary system disorder; osteoporosis; de;
 KW dental disease; dental infection; growth disorder; reproductive disorder;
 KW hypogonadism; fertility disorder; viral infection; bacterial infection;
 KW parasitic infection; metabolic pathway modulation; gene therapy; gene;
 KW zinc metalloprotease; ADAM-TS-7; alpha-2-macroglobulin precursor;
 KW ileal sodium/bile acid cotransporter; prohibitin; MR; C1P4; spinosin;
 KW macrophage stimulating protein precursor; fatty acid-binding protein;
 KW gap junction beta-5 protein; hepsin/plasma transmembrane serine protease.

XX Homo sapiens.

XX WO200233087-A2.

XX 25-APR-2002.

XX 17-OCT-2001; 2001WO-US32496.

XX 17-OCT-2000; 2000US-241040P.

XX 17-OCT-2000; 2000US-241059P.

XX 17-OCT-2000; 2000US-241063P.

XX 17-OCT-2000; 2000US-241243P.

XX 17-OCT-2000; 2000US-241248P.

XX 23-OCT-2000; 2000US-242611P.

XX 23-OCT-2000; 2000US-242612P.

XX 24-OCT-2000; 2000US-242689P.

XX 23-DEC-2000; 2000US-259028P.

XX 20-FEB-2001; 2001US-269613P.

XX 25-APR-2001; 2001US-286324P.

XX 29-MAY-2001; 2001US-294108P.

XX 09-JUL-2001; 2001US-303698P.

XX 16-OCT-2001; 2001US-0981151.

XX (CURA-) CURAGEN CORP.

XX Edinger S, Gerlach V, MacDougall JR, Malyankar UM, Smithson G;

XX Killett J, Peyman JA, Stone BD, Gunther B, Ellerman K, Shimkets RA;

XX Padigaru M, Guo X, Patturajan M, Taupier RJ, Burgess CE; ER;

XX Zernsen BP, Kekula R, Spytek KA, Gangolli EA, Fernandes L;

XX Gorman L;

XX WPI, 2002-590434/63.

XX P-PSDB; ABG76895.

XX Cytoplasmic, nuclear, membrane bound and secreted polypeptides and

XX nucleic acids encoding the polypeptides for diagnosing and treating

XX s.g. cancer, Alzheimer's disease, cardiomyopathy, metabolic disease and

XX diabetes -

XX

XX Claim 8; Page 14-15; 305pp; English.

XX The present invention relates to new NOVX (NOV1-10) polypeptides. The

XX molecules of the invention are useful for treating or preventing a

XX NOVX-associated disorder, such as cardiomyopathy, atherosclerosis, or

XX a disorder related to cell signal processing and metabolic pathway

XX modulation in humans. NOVX polypeptides, nucleic acids and antibodies

XX are useful for treating or preventing disorders or syndromes including

XX breast cancer, Alzheimer's disease, epilepsy, Huntington's disease,

XX anxiety, behavioural disorders, multiple sclerosis, myasthenia gravis,

XX neurodegeneration, Parkinson's disease, pain, stroke, autoimmune

XX disease, allergic, addiction, asthma, endometriosis, graft versus host

XX disease, systemic lupus erythematosus, scleroderma, transplantation,

XX psoriasis, Crohn's disease, HIV (human immunodeficiency virus) infection,

CC atherosclerosis, cirrhosis, rheumatoid arthritis, diabetes,
 CC thrombocytopenia, bleeding disorders, metabolic disorders, obesity,
 CC glucose transport defect, glomerulonephritis, hypercalcaemia, polycystic
 CC kidney disease, pancreatitis, renal tubular acidosis, skin disorders,
 CC congenital diarrhoea, respiratory disease, gastro-intestinal disorders,
 CC muscle, bone, joint and skeletal disorders, haematopoietic disorders,
 CC urinary system disorders, osteoporosis, dental disease and infection,
 CC growth and reproductive disorders, hypogonadism, fertility, and/or other
 CC pathologies and disorders, viral, bacterial, or parasitic infections.
 CC The present nucleic acid sequence encodes a NOVX protein of the
 XX invention.
 XX
 SQ sequence 2432 BP; 567 A; 655 C; 674 G; 536 T; 0 other;
 Query Match 61.6%; Score 1783.6; DB 24; Length 2432;
 Best Local Similarity 81.9%; Pred. No. 0;
 Matches 2374; Conservative 0; Mismatches 34; Indels 490; Gaps 8;
 QY 1 GCTCTCTGATGAAGCCCGCGCGCGCGGATGCGCGGCGCTTGGCGCTTGGATGCTG 60
 DB 22 GCTCTCTGATGAAGCCCGCGCGCGCGGATGCGCGGCGCTTGGCGCTTGGATGCTG 81
 QY 61 TTGCGCGAGGTGGCGAGCAGATGATGTCGCGGCGCTCCGCGAGCGGGAACCGCGG 120
 DB 82 CTGCGCGAGGTGGCGAGCAGATGATGTCGCGGCGCTCCGCGAGCGGGAACCGCGG 141
 QY 121 TC CGGAGCTGCGAGCGAGTGCCTCCGCGGCTCTCTCCGCGAGCGCGCGCTCTCAC 180
 DB 142 TCCGAGCTGCGAGCGAGTGCCTCCGCGGCTCTCTCTCCGCGAGCGCGCGCTCTCAC 201
 QY 181 GCGATGTCGCGCTGTTTTCGCGAGCAGCTGCGGCGCATGGCGCGAGCGCGCGAGC 240
 DB 202 GCGATGTCGCGCTGTTTTCGCGAGCAGCTGCGGCGCATGGCGCGAGCGCGCGAGC 261
 QY 241 GCGTGGGAGCGAGGCTCCGCGCTCTCTCCAGCGCGCGCGCGCGCTGGATGGA 300
 DB 262 GCGTGGGAGCGAGGCTCCGCGCTCTCTCCAGCGCGCGCGCGCGCTGGATGGA 321
 QY 301 AAGCGCGGAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360
 DB 322 AAGCGCGGAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 381
 QY 361 GAAACACCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 417
 DB 382 GAAACACCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 441
 QY 418 GACACAGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 477
 DB 442 GACACAGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 501
 QY 478 GTGCGCTGTGCGAGGTTGAGTCTCTTACACCTTGGCTGAAAGCCCGCAGGCGAGCTTC 537
 DB 502 GTGCGCTGTGCGAGGTTGAGTCTCTTACACCTTGGCTGAAAGCCCGCAGGCGAGCTTC 561
 QY 538 CACATGATGTCAGGAGCTTCCAGCAGCTAGTGGCTCTTGGCTTATTTGCGAGAGTTG 597
 DB 562 CACATGATGTCAGGAGCTTCCAGCAGCTAGTGGCTCTTGGCTTATTTGCGAGAGTTG 621
 QY 598 GGAAGAAGCAGGAGCTAAGTCTGTGAGATTTTACCGCAGAGGAGCTTCTGTGTTCTATCAA 657
 DB 622 GGAAGAAGCAGGAGCTAAGTCTGTGAGATTTTACCGCAGAGGAGCTTCTGTGTTCTATCAA 681
 QY 658 GGCCTTTTGGCATCACAGAAATTCGCATCGATGAGGAGGAGTTCTGTGAGGCTCC 717
 DB 682 GGCCTTTTGGCATCACAGAAATTCGCATCGATGAGGAGGAGTTCTGTGAGGCTCC 741
 QY 718 ACTGCGCTCTGAGCTCTGCAACAGTCAAGAAATGTCCTCCGAGCAGTGTGATCTCCGT 777
 DB 742 ACTGCGCTCTGAGCTCTGCAACAGTCAAGAAATGTCCTCCGAGCAGTGTGATCTCCGT 801
 QY 778 GCTGCTCAGTGTGCGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 837
 DB 802 GCTGCTCAGTGTGCGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 861

QY 838 CCTTACACTCAAGTAGAGACGCCAATTATGCAAACTTACTATGCTATGCGAGAGATTTGAT 897
 DB 862 CCTTACACTCAAGTAGAGACGAGCTATGCAAACTTACTATGCTATGCGAGAGATTTGAT 921
 QY 898 TTCTCTCTTTCTTTGTCATAAATGCAAAAGTGGGACTCCATGCTCGAGAGATAGCCGT 957
 DB 922 TTCTCTCTTTCTTTGTCATAAATGCAAAAGTGGGACTCCATGCTCGAGAGATAGCCGT 981
 QY 958 AATGTTGATAGATGGGATGATGAGCTCAAGTGTGCTGCACATCTCGCAGCATGCC 1017
 DB 982 AATGTTGATAGATGGGATGATGAG-----ATGCC 1014
 QY 1018 CAGCTCCCAAGAGAGACTCTCTCTCTGCGAGAGAGTATAGTTCGTATAGGCGAT 1077
 DB 1015 CAGCTCCCAAGAGAGACTCTCTCTCTGCGAGAGAGTATAGTTCGTATAGGCGAT 1074
 QY 1078 AAGGCTCTCTCTGAGTGTCCATAGAAATGAGAACTGAAGTGGAGACTGGTGGTG 1137
 DB 1075 AAGGCTCTCTCTGAGTGTCCATAGAAATGAGAACTGAAGTGGAGACTGGTGGTG 1134
 QY 1138 GTGAGAAAGAGATGATGAGAACTGAGAACTGAGAAATCAACACTTACGCTCAG 1197
 DB 1135 GTGAGAAAGAGATGATGAGAACTGAGAACTGAGAAATCAACACTTACGCTCAG 1194
 QY 1198 ATACTCAAGAGTATGCTCTTATTCGAAGTGGATGATGCGGAGAGATGGAGCTCGT 1257
 DB 1195 ATACTCAAGAGTATGCTCTTATTCGAAGTGGAGACTGAG-----1238
 QY 1258 CATGACACGCCACTTACTGACTGGCTGGATATATGTTCTCGAGAGATGAGCCCTGT 1317
 DB 1239 -----1238
 QY 1318 GACACTTTGGGATTTGCAACCATAGTGGAAATGTCTAGTAAATATTCGAGCTGACGAT 1377
 DB 1239 -----AGGAAACATCAACATTCGAATTTG 1262
 QY 1378 AATGAGATACAGGCTTGGAGCTGGCTCCACCATGCGCATGCTGAGACACAACTT 1437
 DB 1263 AGGCTGATTC-----1273
 QY 1438 GGCATGATCATGATGGAGAGGAGCATGTGTGTAAGTTCGAGGCAACATCATCTCC 1497
 DB 1274 -----1273
 QY 1498 CCTACATTGGCAGGACGAATGAGTCTCTCTGCTGCTCACCTGCGAGCGCAGATATCTA 1557
 DB 1274 -----1273
 QY 1558 CACAAATTTCAAGCAGCGCTCAAGCTATCTGCTGATGATGAGCAAAAGCTTGTGAG 1617
 DB 1274 -----1273
 QY 1618 GAATACAAATGATCTGAGAAATGCCAGGAAATATATATGATGACAAACACACAGTCCA 1677
 DB 1274 -----1273
 QY 1678 TGCGAGTTCCGAGAGAGAAAGCAAGCTCTGATGCTGGAGCTTTAAAAAGGACATCTGTAA 1737
 DB 1274 -----TTCTAGAGATGAACAGGACATCTGTAA 1302
 QY 1738 GGCCTGTGTGCTCATCTATTGGAAAGGAATGTGAGACTAAATTTTATGCGAGCAGAGAA 1797
 DB 1303 GGCCTGTGTGCTCATCTATTGGAAAGGAATGTGAGACTAAATTTTATGCGAGCAGAGAA 1362
 QY 1798 GGCACAAATTTGTGGGCAATGACATGTGCTGCGGAGGAGACAGTGTGTAATATGTTGAT 1857
 DB 1363 GGCACAAATTTGTGGGCAATGACATGTGCTGCGGAGGAGACAGTGTGTAATATGTTGAT 1422
 QY 1858 GAAGGCCCCCAAGCCACCATGAGGCTGCTGCTGAGCTGCTCTTCTGCTGCCCATGCTCC 1917
 DB 1423 GAAGGCCCCCAAGCCACCATGAGGCTGCTGCTGAGCTGCTCTTCTGCTGCCCATGCTCC 1482

Db	1569	TATATGTCGAAACACACAGTCGAAATGGCGAGCTGTCGAGAGAAAGCCAGCTCTGCATCC	1628
Qy	1712	TGACCTTTAAAGAGACATCTTAAAGCCCTGTGTCCTCATGCTATTTGAGAGAAATCTG	1771
Db	1629	TGACCTTTAAAGAGACATCTTAAAGCCCTGTGTCCTCATGCTATTTGAGAGAAATCTG	1688
Qy	1772	AGACTTAATTTATGCGAGAGCAGAGACCAATTTGTCGGCATGATGATGTCGGCGG	1831
Db	1689	AGACTTAATTTATGCGAGAGCAGAGACCAATTTGTCGGCATGATGATGTCGGCGG	1745
Qy	1832	GAGGACATGTCGTGATATGTCGTGATGAGGGCCCAAGCCCAACCCATGCGCACTGTCGG	1891
Db	1746	GAGGACATGTCGTGATATGTCGTGATGAGGGCCCAAGCCCAACCCATGCGCACTGTCGG	1805
Qy	1892	ACTGTCCTTCTGTCCTCCCATGCTCCAGAGCTTGGCGAGGGGAGTATCTCATAGGAGTC	1951
Db	1806	ACTGTCCTTCTGTCCTCCCATGCTCCAGAGCTTGGCGAGGGGAGTATCTCATAGGAGTC	1865
Qy	1952	GC---CTCTGCACCAACCCCAAGCCATCGATGAGGAGAAATCTGTGAGGGCTCCACTC	2008
Db	1866	GCCTCTCAAAATACACATTTCCAGCCATCGATGAGGAGAAATCTGTGAGGGCTCCACTC	1925
Qy	2009	GCACTCTGAAGCTCTGCAACAGTCAGATAATGTCCTCCGGGACAGTGTGATCTCCGTCGTG	2068
Db	1926	GCACTCTGAAGCTCTGCAACAGTCAGATAATGTCCTCCGGGACAGTGTGATCTCCGTCGTG	1985
Qy	2069	CTCAGTGTGCGGAGCACAAACAGCAGATTCAGAGGGCGGCACTACAGTGGAGGCTT	2128
Db	1986	CTCAGTGTGCGGAGCACAAACAGCAGATTCAGAGGGCGGCACTACAGTGGAGGCTT	2044
Qy	2129	ACACTCAAGTAGAAGATCAGAGCTATATGAACTCTACTGTATCTCGAAGAGGATTCATT	2188
Db	2045	-----GATCAGAGCTTATCGAACTCTACTGTATCTCGAAGAGGATTCATT	2090
Qy	2189	TCTTTCTTTCTTTGTCCTCAATTAAGTCAAAGATGGGACTCCATGCTCGAGGATACCGTA	2248
Db	2031	TCTTTCTTTCTTTGTCCTCAATTAAGTCAAAGATGGGACTCCATGCTCGAGGATACCGTA	2150
Qy	2249	ATGTTTGTATAGATGGGATGTGAGAGAGTCTGATGTGCAATGCTCTTGGATCTGATG	2308
Db	2151	ATGTTTGTATAGATGGGATGTGAGAGAGTCTGATGTGCAATGCTCTTGGATCTGATG	2210
Qy	2309	CTGTTGAGAGCTCTGTGGGTGTGTAACCGGGAATACCTAGCTTGCAACGANTCAAGGG	2368
Db	2211	CTGTTGAGAGCTCTGTGGGTGTGTAACCGGGAATACCTAGCTTGCAACGANTCAAGGG	2270
Qy	2369	GTCTTACACAGCAC	2428
Db	2271	GTCTTACACAGCAC	2312
Qy	2429	GAGCGGAGTTCGGCATCTATGAGAGAGTCTGATGTGCAATGCTCTTGGATCTGATG	2488
Db	2313	GAGCGGAGTTCGGCATCTATGAGAGAGTCTGATGTGCAATGCTCTTGGATCTGATG	2372
Qy	2489	ATGCTCTCAGAGAGTCTACTCAATGGGCACTGGACGCTGAGTGGCGGCGCGGTACA	2548
Db	2373	ATGCTCTCAGAGAGTCTACTCAATGGGCACTGGACGCTGAGTGGCGGCGCGGTACA	2432
Qy	2549	AATTTTCGGGCACTTCTTGGCATCTACAGAGGCTCTTATATGAGCGCGGAGCACTTATCG	2608
Db	2433	AATTTTCGGGCACTTCTTGGCATCTACAGAGGCTCTTATATGAGCGCGGAGCACTTATCG	2492
Qy	2609	CTACTGGACCAACACAGACACTCATTTGTCGATCTGTTTACAGGAGGACACCGGG	2668
Db	2493	CTACTGGACCAACACAGACACTCATTTGTCGATCTGTTTACAGGAGGACACCGGG	2552
Qy	2669	GTGTCCTCGGCACTTCTTGGCATCTACAGAGGCTCTTATATGAGCGCGGAGCACTTATCG	2728
Db	2553	GTGTCCTCGGCACTTCTTGGCATCTACAGAGGCTCTTATATGAGCGCGGAGCACTTATCG	2612
Qy	2729	CCAGCTTACACTTGGGCACTCTGTCGCTGTGAGTCTGCTCCGTCGAGGAGGCTAGGT	2788
Db	2613	CCAGCTTACACTTGGGCACTCTGTCGCTGTGAGTCTGCTCCGTCGAGGAGGCTAGGT	2672

RESULT 12

ABK86137

ID ABK86137 standard; cDNA; 4888 BP.

AC ABK86137;

DT 26-AUG-2002 (first entry)

XX cDNA encoding human protease PRTS4.

Protease; human; gastrointestinal disorder; gastritis; atherosclerosis; ulcerative colitis; Reye's syndrome; cardiovascular disorder; gene; as; hypertension; myocardial infarction; autoimmune disease; AIDS; PRTS; inflammatory disorder; acquired immunodeficiency syndrome; asthma; Grave's disease; cell proliferative disorder; hepatitis; psoriasis; leukaemia; developmental disorder; Cushing's syndrome; impotence; epithelial disorder; dermatitis; scabies; eczema; neurological disorder; Parkinson's disease; dementia; Alzheimer's disease; infertility; Huntington's disease; multiple sclerosis; reproductive disorder.

XX Homo sapiens.

OS Key Location/Qualifiers

1..3711
 CDS /*tag= a
 /product= "PRTS4 protein"

PN W0200238744-A2.

XX 16-MAY-2002.

XX 18-OCT-2001; 2001WO-US51034.

XX 18-OCT-2000; 2000US-241573P.

XX 25-OCT-2000; 2000US-243643P.

XX 02-NOV-2000; 2000US-245256P.

XX 13-NOV-2000; 2000US-248335P.

XX 16-NOV-2000; 2000US-249268P.

XX 20-NOV-2000; 2000US-252303P.

XX 01-DEC-2000; 2000US-250981P.

XX (INCY-) INCYTE GENOMICS INC.

XX Lee BA, Hafalia AJA, Yue H, Lal PG, Yao MG, Lu Y, Wallia NK;

XX Warrick BA, Lu DAW, Baughn MR, Deleane AM, Burford N;

XX Borowsky ML, Lee S, Xu Y, Griffin JA, Kallick DA, Gandhi AR;

XX Arivuch, Ison CH, Tang YT, Azzizai Y, Elliott VS, Swarnakar A;

XX Rankumar J, Nguyen DB, Tribouley CM, Lo TP, Au-young J;

XX Thangavelu K, Kearney L;

XX WPI; 2002-463471/49.

XX P-PSDB; RAU98886.

XX New human proteases useful for diagnosing, preventing or treating

XX anorexia, myocardial infarction, Addison's disease, hepatitis,

XX Cushing's syndrome, eczema, Parkinson's disease, and impotence

XX Claim 5; Page 158-159; 168pp; English.

XX This invention relates to the DNA and protein sequences of novel

XX isolated human proteases (PRTS), the protein sequences of the

XX invention are useful for screening a compound for effectiveness as an

XX agonist or antagonist of its activity, the identified agonist and

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

antagonist are useful for treating a disease or condition associated with decreased or overexpression of functional pRb in a patient. The pRb protein is also useful as an immunogen for preparing polyclonal or monoclonal antibodies by hybridoma technology. An antibody that binds the pRb protein is useful for detection and purification of the protein and can be used to diagnose a condition or disease associated with expression of pRb in a subject or in a biological sample. The sequences of the invention are useful for diagnosis, treatment and prevention of gastrointestinal disorders such as gastritis, ulcerative colitis, Reye's syndrome, etc.; cardiovascular such as atherosclerosis, hypertension, myocardial infarction, etc.; autoimmune/inflammatory disorders such as acquired immunodeficiency syndrome (AIDS), asthma, Grave's disease, etc.; cell proliferative disorders such as Cushing's syndrome, epithelial disorders such as dermatitis, scabies, eczema, etc.; neurological disorders such as Parkinson's disease, dementia, Alzheimer's disease, Huntington's disease, multiple sclerosis, etc.; or reproductive disorders such as infertility, impotence. The present sequence represents the cDNA encoding the human pRb4 protein of the invention.

Query Match 56.6%; Score 1638; DB 24; Length 4888;
Best Local Similarity 93.4%; Pred. No. 0;
Matches 1774; Conservative 0; Mismatches 0; Indels 126; Gaps 1;

QY	1010	ACATGCCAGCCTCCCAAGAGAGACCTTTCATCTTCCAGATGAGTAAAGTCTTGCT	1069
DB	803	ACATGCCAGCCTCCCAAGAGAGACCTTTCATCTTCCAGATGAGTAAAGTCTTGCT	862
QY	1070	TAGGCATTAAGCCTCTCTTCTGAGGTCCCATAGAAATGAAGACTGAAGCTGGAGCT	1129
DB	863	TAGGCATTAAGCCTCTCTTCTGAGGTCCCATAGAAATGAAGACTGAAGCTGGAGCT	922
QY	1130	TGTTGGTGTGCGCAAAAAGATGATGCAAAACCATGGCATGAAATATCACCACTACG	1189
DB	923	TGTTGGTGTGCGCAAAAAGATGATGCAAAACCATGGCATGAAATATCACCACTACG	982
QY	1190	TGCTCAGATATCAACATGGTATCTGCTTTATTCAAAAG	1229
DB	983	TGCTCAGATATCAACATGGTATCTGCTTTATTCAAAAGATGGAACATAGGAGGAACA	1042
QY	1230	-----	1229
DB	1043	TCACATTCGAATTGTAGTCTGTATCTTCTAGAAAGTGAACAGCAGGACTGGTGATA	1102
QY	1230	-----TGGATTGATGGGA	1243
DB	1103	GTACACGAGACACACCTTAAAGTAGTCTTGCAGTGGCAGTCTGGATTGATGGGA	1162
QY	1244	AAGATGGGACTCGTCATGACCGCATCTTACTGACTGGTCTGGATATATGTTCTCGA	1303
DB	1163	AAGATGGGACTCGTCATGACCGCATCTTACTGACTGGTCTGGATATATGTTCTCGA	1222
QY	1304	AGAATGAGCCTGTGACACTTTGGGATTTGACCCATAAGTGGATGTGTAGTAATATC	1363
DB	1223	AGAATGAGCCTGTGACACTTTGGGATTTGACCCATAAGTGGATGTGTAGTAATATC	1282
QY	1364	GCAGTCGACGATTAATGAAGATPACAGGTCTTGGACTGGCTTTCACCAATGCCCATGAT	1423
DB	1283	GCAGTCGACGATTAATGAAGATPACAGGTCTTGGACTGGCTTTCACCAATGCCCATGAT	1342
QY	1424	CTGGACACAACTTTGGCATGATTCATGATGGAGAGGAACATGTGTAAAAGTCCGAGG	1483
DB	1343	CTGGACACAACTTTGGCATGATTCATGATGGAGAGGAACATGTGTAAAAGTCCGAGG	1402
QY	1484	GCACATCATGTCCCTACATTTGGCAGGAGCATGAGTCTTCTCTGGTCAACCTGCA	1543
DB	1403	GCACATCATGTCCCTACATTTGGCAGGAGCATGAGTCTTCTCTGGTCAACCTGCA	1462
QY	1544	GCCTCAGATATCTACACAAATTTCTAAGCACCGCTCAAGCTATCTGCTCTGCTGATCAGC	1603

DB	1463	GCCTCAGATATCTACACAAATTTCTAAGCACCGCTCAAGCTATCTGCTCTGCTGATCAGC	1522
QY	1604	CAAGCCTGTGAAGGATATACAGTATCTCTCAGAAATTCGCCAGAGAAATATATATGACAA	1663
DB	1523	CAAGCCTGTGAAGGATATACAGTATCTCTCAGAAATTCGCCAGAGAAATATATATGACAA	1582
QY	1664	ACACACATGCTCAAGTGGCAGTTCGAGAGAAAGCCACAGCTTCGCATCTGGACTTAAAA	1723
DB	1583	ACACACATGCTCAAGTGGCAGTTCGAGAGAAAGCCACAGCTTCGCATCTGGACTTAAAA	1642
QY	1724	AGCAGCTGTAAAGCCCTGTGTCCTCATCTGATATGGAAGAAATGAGACTTAAATTA	1783
DB	1643	AGCAGCTGTAAAGCCCTGTGTCCTCATCTGATATGGAAGAAATGAGACTTAAATTA	1702
QY	1784	TGCCAGCAGCAGGACCAATTTGTCGCGATGACATGCTGGTGGGAGGACAGTGTG	1843
DB	1703	TGCCAGCAGCAGGACCAATTTGTCGCGATGACATGCTGGTGGGAGGACAGTGTG	1762
QY	1844	TCARAATGTGATGAGGCCCCACAGGCCCATCTGCGCACTGGTCTGCTCTCTT	1903
DB	1763	TGARAATGTGATGAGGCCCCACAGGCCCATCTGCGCACTGGTCTGCTCTCTT	1822
QY	1904	GGTCCCCATGCTCCAGGACCTGCGGGGGGAGTATCTCATAGGAGTCCCTCTGCACCA	1963
DB	1823	GGTCCCCATGCTCCAGGACCTGCGGGGGGAGTATCTCATAGGAGTCCCTCTGCACCA	1882
QY	1964	ACCCACAGCATTCGATGAGGAGGAGTCTGTGAGGGCTCCACTGCACTTGGAGCTCT	2023
DB	1883	ACCCACAGCATTCGATGAGGAGGAGTCTGTGAGGGCTCCACTGCACTTGGAGCTCT	1942
QY	2024	GCACAGTCAGAAAATGTCCTCCCGGACAGTGTGTGATTCGCTCGCTGCTGCGGAGC	2083
DB	1943	GCACAGTCAGAAAATGTCCTCCCGGACAGTGTGTGATTCGCTCGCTGCTGCGGAGC	2002
QY	2084	ACACACACAGCATTCAGAGGGGGGACATCAAGTGGAGCTTACACTCAAGTAGAG	2143
DB	2003	ACACACACAGCATTCAGAGGGGGGACATCAAGTGGAGCTTACACTCAAGTAGAG	2062
QY	2144	ATCGAGCTATGCAAACTCTACTGTGCGAGAGGATTAATCTCTCTCTCTCTTGT	2203
DB	2063	ATCGAGCTATGCAAACTCTACTGTGCGAGAGGATTAATCTCTCTCTCTCTTGT	2122
QY	2204	CAATTAAGTCAAGATGGGACTCTGCTGCGAGGATAGCGTATGTTGTATAGATG	2263
DB	2123	CAATTAAGTCAAGATGGGACTCTGCTGCGAGGATAGCGTATGTTGTATAGATG	2182
QY	2264	GGATATGCGAGAGTGGGATGATGATGCTGCGAGGATAGCGTATGTTGTATAGATG	2323
DB	2183	GGATATGCGAGAGTGGGATGATGATGCTGCGAGGATAGCGTATGTTGTATAGATG	2242
QY	2324	GTGGGTTGTAAACGGGATTAATCAATGCTGCAAGTCTGCGAGGCTCTACACAGC	2383
DB	2243	GTGGGTTGTAAACGGGATTAATCAATGCTGCAAGTCTGCGAGGCTCTACACAGC	2302
QY	2384	ACCAACACACACACACACATGATGATGATGATGATGATGATGATGATGATGATGATG	2443
DB	2303	ACCAACACACACACACACATGATGATGATGATGATGATGATGATGATGATGATGATG	2362
QY	2444	GCATCTATGAAATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2503
DB	2363	GCATCTATGAAATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2422
QY	2504	ACTACTGATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2563
DB	2423	ACTACTGATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2482
QY	2564	CTTTCCAGCTACAGAGCTGCTTATGAGGCTGCTTATGAGGCTGCTTATGAGGCTG	2623
DB	2483	CTTTCCAGCTACAGAGCTGCTTATGAGGCTGCTTATGAGGCTGCTTATGAGGCTG	2542
QY	2624	ACGAGACATGATGAGGCTGCTTATGAGGCTGCTTATGAGGCTGCTTATGAGGCTG	2683
DB	2543	ACGAGACATGATGAGGCTGCTTATGAGGCTGCTTATGAGGCTGCTTATGAGGCTG	2602

QY 2684 ACTCCATGCTGGTGGGACCGAGACAGCCCTCCACGCCAGCTACACTGGG 2743
 DB 2603 ACTCCATGCTGGTGGGACCGAGAGAGCCCTCCACGCCAGCTACACTGGG 2662
 QY 2744 CCATCGTGGCTCTGAGTCTCCCTGCTCTGGGAGGGG 2783
 DB 2663 CCATCGTGGCTCTGAGTCTCCCTGCTCTGGGAGGGG 2702
 RESULT 13
 ID ABK49821 standard; cDNA; 4234 BP.
 XX
 AC ABK49821;
 XX
 DT 15-JUL-2002 (first entry)
 XX
 DE Human cDNA encoding partial ADAMTS protein #1.
 XX
 KW Human; ss; gene; ADAMTS; cytostatic; antidiabetic; antirheumatic;
 KW antiarthritic; aneurysm; vulnery; neovascularisation; angiodysplasia;
 KW diabetic omentopathy; chronic rheumatoid arthritis; gene therapy;
 KW refractory skin ulcer; gastric ulcer; post-operative healing failure;
 KW repolysin-type 2N-metalloproteinase domain; disintegrin-like domain; TSP1;
 KW thrombospondin type 1 domain; sexual cycle; tumour; 5p-syndrome deletion;
 KW chromosome 5p15.2-15.3; Cri-du-chat syndrome.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 CDS 3..3068
 FT /tag= a
 FT /product= "ADAMTS"
 FT /partial
 FT /note= "No start codon"
 XT
 W0200231163-Al.
 XX
 XX 18-APR-2002.
 XX
 XX
 XX 11-OCT-2001; 2001WO-JP08913.
 XX
 XX 11-OCT-2000; 2000JP-0311309.
 XX 02-APR-2001; 2001JP-0102905.
 XX
 XX (KAZU-) KAZUSA DNA RES INST FOUND.
 XX (MITS-) MITSUBISHI PHARMA CORP.
 XX
 XX Ohara O, Nagase T, Nomura N, Yano K, Murakami K, Yasuda S;
 XX Kanazaki K;
 XX
 XX WPI: 2002-372277/40.
 XX P-PSDB; BAU79496.
 XX
 XX Human brain-originated ADAMTS family polypeptide and encoded gene;
 XX applicable in diagnosis and screening compounds for drug compositions
 XX in treating diseases due to e.g. neovascularisation
 XX
 XX Claim 4; Page 127-135; 172pp; Japanese.
 XX
 XX The invention relates to a polypeptide belonging to the ADAMTS family, is
 XX selected from sequences appearing as BAU79496, BAU79497 and BAU79499,
 XX a protein that contains the amino acid sequence of the polypeptide, or a
 XX 50% homology with the amino acid sequence of the polypeptides or a
 XX polypeptide modified from any of the polypeptides but with some amino
 XX acids deleted, substituted, added or inserted also included in the
 XX polynucleotides encoding the polypeptides (or their complementary
 XX strands or variants), a recombinant vector containing any of the
 XX polynucleotides, a transformant which is transformed with the recombinant
 XX vector, producing the polypeptide, protein or peptide by culturing the
 XX transformant, an antibody that can recognize the polypeptide, protein or
 XX peptide and screening compounds to promote or inhibit activity of the

CC polypeptide or protein, or to promote or inhibit expression of the
 CC polynucleotide by using the polypeptide, protein, peptide, the
 CC polynucleotide, vector, transformant or/and antibody, particularly in
 CC the presence of a test compound for contact before evaluating the
 CC activity by measuring signal changes. The polypeptide and encoded gene
 CC are applicable in diagnosis and screening compounds for drug compositions
 CC in treating diseases due to neovascularisation, diabetic omentopathy,
 CC chronic rheumatoid arthritis, angiodysplasia, refractory skin and gastric ulcers
 CC and post-operative healing failure, including gene therapy.
 CC The gene encoding such polypeptide has conserved repolysin-type 2N-
 CC metalloproteinase domain, disintegrin-like domain and TSP1 (thrombospondin
 CC type 1) domain. Its encoded protein is characterised by high expression
 CC in ovaries, changes in expression dose depending on the sexual cycle, a
 CC decrease in tumour cell and location of the gene on the 5p-syndrome
 CC deletion site on chromosome 5p15.2-15.3 (associated with
 CC Cri-du-chat syndrome). The present sequence encodes a ADAMTS
 CC protein of the invention.
 XX
 SQ Sequence 4234 BP; 1036 A; 1103 C; 1130 G; 965 T; 0 other;
 Query Match 56.5%; Score 1634.8; DB 24; Length 4234;
 Best Local Similarity 93.3%; Pred. No. 0;
 Matches 1772; Conservative 0; Mismatches 2; Indels 126; Gaps 1;
 QY 1010 ACATGCCCGAGCTCCCAAGAGAGACCTCTTCATCTTCCGAGATGATTAAGTCTTGT 1069
 DB 157 ACATGCCCGAGCTCCCAAGAGAGACCTCTTCATCTTCCGAGATGATTAAGTCTTGT 216
 QY 1070 TACGCGATAGCCCTCTCTCTGAGTCCCATAGAAATGAAAGACTGAACCTGGAGACCT 1129
 DB 217 TACGCGATAGCCCTCTCTCTGAGTACCATAGAAATGAAAGACTGAACCTGGAGACCT 276
 QY 1130 TGGTGGTGGTGGACAAAAAGATGATGCAAAACACATGCGCATGAAAAATATCACTACG 1189
 DB 277 TGGTGGTGGTGGACAAAAAGATGATGCAAAACACATGCGCATGAAAAATATCACTACG 336
 QY 1190 TGCTTACGATCTCAACATGGTATCTGCTTTATTCAAAGA 1229
 DB 337 TGCTTACGATCTCAACATGGTATCTGCTTTATTCAAAGA 396
 QY 1230 ----- 1229
 DB 397 TCAACATTGCAATTGTAGGTCTGATTTCTTTAGAAATGAAAGACGAGCATGCTGATAA 456
 QY 1230 -----TGATTGATGGGA 1243
 DB 457 GTCACCGCAGACACACACTTAAGTAGTCTTGCAGCTGGCAGTCTGATTTGATGGGA 516
 QY 1244 AAGATGGGAGCTGCTATGACCCGATCTTACTGATCTGCTGGATATATGTTCTGGA 1303
 DB 517 AAGATGGGAGCTGCTATGACCCGATCTTACTGATCTGCTGGATATATGTTCTGGA 576
 QY 1304 AAGATGGGAGCTGCTGACCTTTGGGATTTGCACTCCATAGTGGATGTTAGTAAATATC 1363
 DB 577 AAGATGGGAGCTGCTGACCTTTGGGATTTGCACTCCATAGTGGATGTTAGTAAATATC 636
 QY 1364 GCGAGTGCACGATTAATAGATACATAGTCTTGGATCTGGCTTTCACCATTCGCCATGAGT 1423
 DB 637 GCGAGTGCACGATTAATAGATACATAGTCTTGGATCTGGCTTTCACCATTCGCCATGAGT 696
 QY 1424 CTGCGACACACTTTGGCATGATTCATGATGGAGAGGGAACATGTTGTAAGTCCGAGG 1483
 DB 697 CTGCGACACACTTTGGCATGATTCATGATGGAGAGGGAACATGTTGTAAGTCCGAGG 756
 QY 1484 GCGACATCATGCTCCCTTCATTCGCGAGGAGGCAATGGAGTCTTCTCTGCTGGTCACTGCA 1543
 DB 757 GCGACATCATGCTCCCTTCATTCGCGAGGAGGCAATGGAGTCTTCTCTGCTGGTCACTGCA 816
 QY 1544 GCGCGCATGATCTACACAAATTTCTAAGACCGCTCAAGCTATCTGCTCTGATGATGAC 1603
 DB 817 GCGCGCATGATCTACACAAATTTCTAAGACCGCTCAAGCTATCTGCTCTGATGATGAC 876
 QY 1604 CAAGCGCTGGAGGGAATACATGATCTCTGAGAAATTCGCGAGGGAATATATGATGCA 1663

DB 877 CAAAGCCCTGAGGAGATACAGATATCTGAGAAATGCGAGAGAAATATATGATGCA 936
QY 1664 ACACACAGTGCAGTGCAGTTCGAGAGAGAGAGCCAGCTCTGATCTGCGATTTAA 1723
DB 937 ACACACAGTGCAGTGCAGTTCGAGAGAGAGAGCCAGCTCTGATCTGCGATTTAA 996
QY 1724 AGGACATCTGTAAGACCTGTGTGGCCATGTAATGGAAGAAATGTGAGACATAAATTA 1783
DB 997 AGGACATCTGTAAGACCTGTGTGGCCATGTAATGGAAGAAATGTGAGACATAAATTA 1056
QY 1784 TCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1843
DB 1057 TCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1116
QY 1844 TGAATATGATGATGAGAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCA 1903
DB 1117 TGAATATGATGATGAGAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCA 1176
QY 1904 GGTCCCATGCTCCAGAGCTCCGAGAGAGAGAGATCTCATAGAGAGTCCCTCGCACCA 1963
DB 1177 GGTCCCATGCTCCAGAGCTCCGAGAGAGAGAGATCTCATAGAGAGTCCCTCGCACCA 1236
QY 1964 ACCCAGGACATCGCATGGAGGAGATCTGTGAGGAGTCCACTCGCATCTGAGAGCTCT 2023
DB 1237 ACCCAGGACATCGCATGGAGGAGATCTGTGAGGAGTCCACTCGCATCTGAGAGCTCT 1296
QY 2024 GCACAGCTCAGAAATGTCCTCCGAGAGAGTGTGATCTTCCTGCTGCTGCTGCTGCTG 2083
DB 1297 GCACAGCTCAGAAATGTCCTCCGAGAGAGTGTGATCTTCCTGCTGCTGCTGCTGCTG 1356
QY 2084 ACACACAGCAGCAGTCCAGAGGCGGCACTACAGCTGAGAGCTTACACTCAAGTAGAG 2143
DB 1357 ACACACAGCAGCAGTCCAGAGGCGGCACTACAGCTGAGAGCTTACACTCAAGTAGAG 1416
QY 2144 ATCAGGACTTACAGAACTCTACTGATCGAGAGAGATTTGATTTCTCTCTCTCTCT 2203
DB 1417 ATCAGGACTTACAGAACTCTACTGATCGAGAGAGATTTGATTTCTCTCTCTCTCT 1476
QY 2204 CAATAAAGTCAAAATGGAGCTCCATGCTCCGAGAGATGCGCTAAATTTCTGATAGT 2263
DB 1477 CAATAAAGTCAAAATGGAGCTCCATGCTCCGAGAGATGCGCTAAATTTCTGATAGT 1536
QY 2264 GGNATATGAGAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2323
DB 1537 GGNATATGAGAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1596
QY 2324 GTGGGTTGTAAAGGAGTAATCACTCAGCTGCAAGATTTCAAGAGGCTCTTACACCAAG 2383
DB 1597 GTGGGTTGTAAAGGAGTAATCACTCAGCTGCAAGATTTCAAGAGGCTCTTACACCAAG 1656
QY 2384 ACCACACACCAACAGTATATACATGATGATGATGATGATGATGATGATGATGATG 2443
DB 1657 ACCACACACCAACAGTATATACATGATGATGATGATGATGATGATGATGATGATG 1716
QY 2444 GCATCTATGAATGAAGTCTTACCTCTTACATTTCTGCGCAATGCGCTCAGAGGT 2503
DB 1717 GCATCTATGAATGAAGTCTTACCTCTTACATTTCTGCGCAATGCGCTCAGAGGT 1776
QY 2504 ACTACCTGATGAGCAGTGCAGCTGATGATGATGATGATGATGATGATGATGATGATG 2563
DB 1777 ACTACCTGATGAGCAGTGCAGCTGATGATGATGATGATGATGATGATGATGATGATG 1836
QY 2564 CTTTCGACTACAGAGCTCTATATAGCAGCAGGAGTATGCTGCTGCTGCTGCTGCTGCTG 2623
DB 1837 CTTTCGACTACAGAGCTCTATATAGCAGCAGGAGTATGCTGCTGCTGCTGCTGCTGCTG 1896
QY 2624 ACGAGACACTGATTTGAGAGCTGCTGTTTCAGGAGAGAGACCGGCTGCTGCTGCTGCTG 2683
DB 1897 ACGAGACACTGATTTGAGAGCTGCTGTTTCAGGAGAGAGACCGGCTGCTGCTGCTGCTG 1956
QY 2684 ACTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2743

DB 1957 ACTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2016
QY 2744 CCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2783
DB 2017 CCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2056

RESULT 14
ABK4927
ID ABK4927 standard, DNA, 7660 BP.
XX ABK4927;
AC ABK4927;
XX
DT 15-JUL-2002 (first entry)
XX
DE Plasmid pFastBac1-MS/HI-F001256-1.
KW ds; ADAMTS, cystostatic; antidiabetic; antirheumatic; cyclic;
KW antiarthritis; antiulcer; vulnerable; neovascularisation; angina;
KW diabetic osteomyelitis; chronic rheumatoid arthritis; gene therapy;
KW erythrocytic skin ulcer; gastric ulcer; post-operative healing failure;
KW repressin-type 2H-metalloproteinase domain; disintegrin-like domain; TSP1;
KW chromosome 5p15.3; 3p21.3; C11-du-chat syndrome; circular;
KW chromosome 5p15.3; 3p21.3; C11-du-chat syndrome; circular;
KW pFastBac1-MS/HI-F001256-1.
XX Homo sapiens.
OS Escherichia coli.
OS Baculovirus.
OS Rhesus macaque polyoma virus.
OS Transposon f17.
XX
XX
PW W020231163-A1.
XX
XX 18-APR-2002.
XD
XX
PF 11-OCT-2001; 2001WO-JP08913.
XX
PR 11-OCT-2000; 2000JP-0311309.
PR 02-APR-2001; 2001JP-0102905.
XX
XX (KAZU-) KAZUSA DNA RES INST FOUND.
PA (MITS-) MITSUBISHI PHARMA CORP.
XX
PI Ohara O, Negase T, Nomura N, Yano K, Murakami K, Yasuda S;
PI Kanzaki K,
XX
XX WPI; 2002-372277/40.
XX
PT Human brain-originated ADAMTS family polypeptide and encoded gene,
PT applicable in diagnosis and screening compounds for drug compositions
PT in treating diseases due to e.g. neovascularisation -
PS Example 20; Page 163-167; 172pp; Japanese.
XX
CC The invention relates to a polypeptide belonging to the ADAMTS family is
CC selected from sequences appearing as AAU79496, AAU79497 and AAU79499,
CC a protein that contains the polypeptide, a protein having not less than
CC 50% homology with the amino acid sequence of the polypeptides or a
CC polypeptide modified from any of the polypeptides but with some amino
CC acids deleted, substituted, added or inserted. Also included are the
CC polynucleotides encoding the polypeptides (or their complementary
CC strands or variants), a recombinant vector containing any of the
CC polynucleotides, a transformant which is transformed with the recombinant
CC vector, producing the polypeptide, protein or peptide by culturing the
CC transformant, an antibody that can recognize the polypeptide, protein or
CC peptide and screening compounds to promote or inhibit activity of the
CC polypeptide or protein, or to promote or inhibit expression of the
CC polynucleotide by using the polypeptide, protein, peptide, the
CC polynucleotide, vector, transformant or/and antibody, particularly in
CC the presence of a test compound for contact before evaluating the
CC activity by measuring signal changes. The polypeptide and encoded gene

CC are applicable in diagnosis and screening compounds for drug compositions
CC in treating diseases due to neovascularisation, diabetic omentopathy,
CC chronic rheumatoid arthritis, angiona, refractory skin and gastric ulcers
CC and post-operative healing failure, including gene therapy.

CC The gene encoding such polypeptide has conserved repressi

Query Match 54.5%; Score 1578.4; DB 24; Length 7668;
Best Local Similarity 93.1%; Pred. No. 0;
Matches 1715; Conservative 0; Mismatches 1; Indels 126; Gaps 1;

Qy	1068	CTTACGGCATTAAGCGCTCTCTCTGAGGTCCTCATAGAAATGAGAACTCGAACTCGAGC	1129
Db	4106	CTTACGGCATTAAGCGCTCTCTCTGAGGTCCTCATAGAAATGAGAACTCGAACTCGAGC	4165
Qy	1128	CTTGGCTGGTGTCGACAAAGAGATGCAAAACCAATGCGCATGAAATATCACCACTTA	1187
Db	4166	CTTGGCTGGTGTCGACAAAGAGATGATGAAAACCAATGCGCATGAAATATCACCACTTA	4225
Qy	1188	GGTGGCTCAGCATCTCAACATGATATCTCTCTTATTCBAAG	1229
Db	4226	GGTGGCTCAGCATCTCAACATGATATCTCTCTTATTCBAAGTGTACAAATAGGAA	4285
Qy	1230	-----	1229
Db	4286	CATCAACATTGCAATTGAGTCTGATTCTCTCTAGAAGATGACAGCCAGGACTGTGAT	4345
Qy	1230	-----	1241
Db	4346	AAGTCACCACGACGACACACACTTAAAGTAGCTCTGCGCAGTGGCAGCTCTGGATTGATGG	4405
Qy	1242	GAAAGATGGGACTCGTCATGACACGCGCATCTTACTGACTGCTCTGATATATGTCCTG	1301
Db	4406	GAAAGATGGGACTCGTCATGACACGCGCATCTTACTGACTGCTCTGATATATGTCCTG	4465
Qy	1302	GAGAAATGAGCCCTGTGACACTTTTGGGATTTGCACCCATTAAGTGGAAATGTGTAGTAAATA	1361
Db	4466	GAGAAATGAGCCCTGTGACACTTTTGGGATTTGCACCCATTAAGTGGAAATGTGTAGTAAATA	4525
Qy	1362	TGCGAGCTCCACGATTAATGAAGATACAGTCTTGTGACACTGGCCCTTCAACATTGCCCATCA	1421
Db	4526	TGCGAGCTCCACGATTAATGAAGATACAGTCTTGTGACACTGGCCCTTCAACATTGCCCATCA	4585
Qy	1422	GTCTGGACACAACTTTGGCATGATTCAATGATGGAGAGGGAACATGTGTAAAAAGTCCGA	1481
Db	4586	GTCTGGACACAACTTTGGCATGATTCAATGATGGAGAGGGAACATGTGTAAAAAGTCCGA	4645
Qy	1482	GGGCAACATCATGTGCCCTTACATTTGGCAGGACGCAATGGAGTCTTCTCTGGTCAACCCGT	1541
Db	4646	GGGCAACATCATGTGCCCTTACATTTGGCAGGACGCAATGGAGTCTTCTCTGGTCAACCCGT	4705
Qy	1542	CAGCCGCCAGTATCTACAAATTTCTAAGCACCGCTCAAGCTATCTGGCTTGTCTGATCA	1601
Db	4706	CAGCCGCCAGTATCTACAAATTTCTAAGCACCGCTCAAGCTATCTGGCTTGTCTGATCA	4765
Qy	1602	GCCAAAGCCCTGTGAAGGAATACAAGTATCCTTGAGAAATTTGCCAGGAGAAATATATGATGC	1661
Db	4766	GCCAAAGCCCTGTGAAGGAATACAAGTATCCTTGAGAAATTTGCCAGGAGAAATATATGATGC	4825
Qy	1662	AAACACACAGTGGAGTGGCAGTTCGAGAGAGAAAGCCAAAGCTCTGCATGCTGGACTTTAA	1721
Db	4826	AAACACACAGTGGAGTGGCAGTTCGAGAGAGAAAGCCAAAGCTCTGCATGCTGGACTTTAA	4885
Qy	1722	AAAGGACATCTGTAAAGCCCTGTGGTGCAATCGTATTTGAAGGAATTTGTGAGACTAAAT	1781

[illegible]

QY	1418	ATGAGTCTGGACCAACATTTTGGCATCATGATTCGAGAGGGAACATGTGTAAAAAGT	1417
DB	1782	ATGAGTCAAGGCGACAACTTTTGGCATCGTTATCATGATCGTGAAGCGAATCCCTTCGACGAAGA	1841
QY	1478	CCGAGGGCAACATCATGTCCCTCATATTGGCAGGAGCGCATGGAGTCTTCTTCCCTGGTCAC	1537
DB	1842	CAGAAAGGCAACATCATGTCAATCCACCACTGATCTGAAACAAATGAGGGGTGTTTTCATGGTGGT	1901
QY	1538	CTTCGAGCGCGCGAGTATCTACACAAATTTCTAAGCAGCGCTCAAGCTATCTGGCTTGGCTG	1597
DB	1902	CTCTCAGCCGACGATCTCTAAAGAAATTCCTCTAGTACACCGAGGCTGGCTGTCTGGTGGG	1961
QY	1598	ATCAGCCAAAGCGCTGTGAAGGAATACAGAGTATCTCGAGAAATTCGCCAGGGAATATATATG	1657
DB	1962	ATGAGGCCAAGCAAAAGGGACAGATATAAATATCCGACAAACATCCCGGGGACGATTTATG	2021
QY	1658	ATGCAAAACACACAGTCGAATGTGGCAGTTCCGAGAGAAAGCCAACTCTGCATCTGGACT	1717
DB	2022	ATGCCGACATGCAATGTAAAGTGGCAATTTTGAGCGCAAGCCAAAGCTGTGGAGCTTGGGG	2086
QY	1718	TTAAAAAGACACTCTGTAAAGCCCTCTGTGTGCCATGTATTTGGAGAAATATGTGAGACTA	1777
DB	2082	TTATGAAGAGCAATTTGCAAGTCACTCTGTGTGCCACGCGGTGGGCCACAGATGTGAAACCA	2141
QY	1778	AAATTTATGCCACGACGAGCAAGGSCAAATTTTGGGGCATGACATGTGTGCCCGGGAGAC	1837
DB	2142	AGTTATGCTGCTCAGCAGCAGAGGGACCGCTGTGTGCTTGAGTATGTGTGTCTGGCAGAGCC	2201
QY	1838	AGTGTGTAAATATGTGTGATGAGAGGCCCCCAAGCCACCAATGGCCACTGCTCGACATGGT	1897
DB	2202	AGTGTGTAAAGCTCGGAGAACTATGGGCCCCCGGCCCAATCCGCGCAAGTGTCTGCTGGT	2261
QY	1898	CTTCTGTGCCCATGCTTCAGACACTCTGGAGGGGAGCTATCTCATAGGAGTGGCTCT	1957
DB	2262	CGAGTGGTCCAGATGTTTCTGCACCTCTGGCGAGGGGTCAAGTTCAGAGAGCACT	2321
QY	1958	GCACCAACCCGAGGCATCGATGAGGAGGATCTGTGGAGGCTCCACTCGCACTCTGA	2017
DB	2322	GTAGTAAACCCAGCGCTCAGTATGTGCGAGTACTGTCCAGAACTAGCGCACTCTCA	2381
QY	2018	AGCTCTGCAACAGTCAGAAATGTGTCCCGGACAGTGTTCACCTCCGTGCTCGTGGTGG	2077
DB	2382	AGCTGTCAACATTAACCTTGTCTCGAGATAGCTGGATTTCCGTGCCCCCAATGATGTG	2441
QY	2078	CGCGACACACGACGACGATTCAGAGGGCGCATACAAGTGGAGCCCTACACTCAAG	2137
DB	2442	CAGAGTATACACAGCCCTTCCCTGGATGGTGTACCTTTGGAACCTCTACACAAAG	2501
QY	2138	TAGAGATCAGACCTTATGCAAACTCTACTGTATTCGACAGAGATTTGATTTCTCTCTTT	2197
DB	2502	TTGAGAGGAAGATCGATGTAAACCTCTACTGCAAGGCGGAGAACTTTGAGTTTCTCTTG	2561
QY	2198	CTTTGTCAATAAAGTCAAAAGATGSGACTCCATGCTCGGAGGATAGCCGATATGTGTGA	2257
DB	2562	CCATGCTCGCAAGGTGAAAGATGGAGCGCTTGTTCGCCACAGAAATGACCTCTGCA	2621
QY	2258	TAGATGGGATGTGACAGAGTTGGATGTGCAATGTCTCTTGGATCTGATGCTGTGGAG	2317
DB	2622	TCGATGGGATTTGTGAACCTGGTGGAGTGTGACATAGAGCTTGGCTCTTAAGSCGGTTC	2681
QY	2318	ACGTCTGTGGGGTGTGTAAACGGGAATAACTCAGCTCGACAGATTCACAGGGGTCTCTACA	2377
DB	2682	ATGCATGTGGTGTGTCNAAGCGGATTAACCTCAACTCTCAAGTTTATAAGGSCCTGTACC	2741
QY	2378	CAAGCCACCAACACACCAACAGTATTATCAATGGTCACCACTCTCTTCTGGAGSCCGGA	2437
DB	2742	TCAGTCACCAAAAGAAATGAGTATTACCCCTGTGGTCAACATCCCGCTGGGGCCCGAA	2801
QY	2438	GTATCCGCACTATCAAAATGACGTCCTACTCTCTCAATTTCTGTGCGCAATGCCCTCA	2497
DB	2802	GCATTTGAGATTCAGAGGTTGAGCTTTCTTCAGCTATCTCTGCTGTTCGAGCCCTCACTG	2861

QY	2498	GAAGCTACTCTCGTAATGSGCACTGCACCGTGGGACTGGCCGGCGGTACAAAATTTTCGG	2553	
DB	2862	AAAGTAGTATTACCTCACAGGSGGGTGGAGCATCGACTGGCCAGGSGGACTTCACCTTGGCGAS	2921	
QY	2558	GCACCTACTTTTCGACTACAGCGGTCTCTATATGAGCCGAGAACTTAATCGTCTACTGAGAC	2617	
DB	2922	GGACACGTTTGAATACACAGCGTTCCTTTTAAACGCGCTTGAACGACTGTATGACACAGGAC	2981	
QY	2618	CAACCAACGAGCACTGATTGTGGAGCTGCTGTTTTCAGGGAAGAAACCCGGGTGTTCCT	2677	
DB	2982	CCACGAATAGACGCTGGTCTTGTGAATTTCTGACACAGGSCAGAAATCCGGGGATCGCTT	3041	
QY	2678	GGGAATACTCTCATCTGCTTGGGACCGCAGAAAGCAGCCCTTGGCCAG---CCGAGCT	2734	
DB	3042	GGAAAGTATGCACTTTCCCAAGGTGATGAATTAATCTCAGCGAGCCACCAAGAGATATCACCC	3101	
QY	2735	ACACTGTGGGCGCATCGTGGCGCTCTCGAGTGTCTCGGTCTCTGGGAGGAGGGCT	2784	
DB	3102	ACACCTGGGCGACAGTGCAGTCTGCACTGTCTCAGTTCACCTCGGCTGGAGGT	3151	
RESULT 3	AK036137	4593 bp	linear	HTC 05-DEC-2002
LOCUS	Mus musculus	16 days neonate cerebellum cDNA, RIKEN full-length enriched library, clone:9630038L21 product:hypothetical Metalloprotease (ADAM type),prelysin (M12B) family containing protein, full insert sequence.		
ACCESSION	AK036137	GI:26085046		
VERSION	AK036137.1			
KEYWORDS	HTC; CAP trapper.			
SOURCE	Mus musculus (house mouse)			
ORGANISM	Mus musculus			
REFERENCES	1	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus; Mus.		
AUTHORS	Carninci, P. and Hayashizaki, Y.			
TITLE	High-efficiency full-length cDNA cloning			
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)			
MEDLINE	99279253			
PUBMED	10349636			
REFERENCES	2	Kanno, H., Shibata, Y., Hayatsu, N., Sugawara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes			
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)			
MEDLINE	20499374			
PUBMED	11042159			
REFERENCES	3	Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carninci, P., Kanno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.		
TITLE	RIKEN integrated system analysis (RISA) system--384-format sequencing pipeline with 384 multipillar sequencer			
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)			
MEDLINE	20530913			
PUBMED	11076851			

REFERENCES

Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, K., Fukuishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, T., Hara, A., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojohori, T., Bono, H., Kasukawa, T., Saito, E., Kadota, K., Matsuda, H., Ashbourn, M., Batalov, S., Casavent, T., Plesschmann, W., Gaasterland, T., Giscl, C., King, B., Kochiwa, T., Nishikawa, T., Kato, Y., Nishigaki, N., Nakagawa, S., Tomita, M., Nishikawa, Y., Schram, M., Staubli, F., Suzuki, R., Tanaka, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Bareth, G., Blake, J., Boffelli, D., Botlung, N.,

Carinci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Flechner, C., Fujita, M., Gariboldi, M., Gustinich, S., Hill, D., Hofmann, J., Hume, D.A., Kany, M., Lee, N.H., Lyons, P., Nordone, P., Perchomil, I., Rasmussen, R., Raza, S., Shihata, Y., Storch, K.F., Sung, E., Tschopp, C., Wenz, C., Shihata, Y., Storch, K.F., Suzuki, H., Teraoka, K., Wang, K.H., Weitz, C., Whitaker, C., Wilming, L., Wong, K., Koyas, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohzuki, S., and Hayashizaki, Y.	QY	1296	CTGGAGAGATGAGCCCTGTGACACTTGGATTTGGACCCCTAGAGTGGATGTGACTA	1357
Nature 409 (6821), 685-690 (2001)	DB	788	CTGGAGAGACGACGACGATGTGACACTAGATTTGCTTCTATGATGGATGTGACTA	847
11217851	QY	1358	AATATGCGCAGCTGACAGATTAATGAAGATACAGTCTTGGACTGSCCTTCCACATGSCC	1417
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.	DB	848	AGTACCGAAGCTGTACCTCAATGAGACACAGSCTTGGCTTGGCTTCCACAAATGCGC	907
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	QY	1418	ATGAGTCTGACACACAACTTTGGCATGATTCATGATGGAGAGGGAACATGTGTAAGAAGT	1477
Nature 420, 563-573 (2002)	DB	908	ATGAGTCAAGGCGACAACTTTGGCATGATTCATGATGGTGAAGGCAATCCCTCGAGAAAG	967
6 (bases 1 to 4593)	QY	1478	CCGAGGCGACATCATGTCCTTACATGCGAGACGCAATGAGTCTTCTCTGCTGAC	1537
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanezaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihar, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Shihata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akashi, S., Takeda, Y., Tanaka, T., Tomaru, A., Toyota, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.	DB	968	CAGAGGCGACATCATGTCACCCACACACTGACTGGAACAAATGGGTGTTCATGCTGCT	1027
Submitted (16-JUL-2001) Yoshhide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Sukenri-cho, Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp, URL: http://genome.res.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)	QY	1538	CCTGACAGCCCGCAGTATCTACACAAATTTCTAAGCACCCTCAAGCTATCTGCTTGTCTG	1597
CNA library was prepared and sequenced in Mouse Genome Encyclopedia project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse cDNAs.	DB	1028	CCTGACGCGACAGTATCTAAGAAATTCCT	1058
URL: http://genome.res.riken.go.jp/	QY	1598	ATCAGCCAAAGCCTGTGAAGGAATACAGTATCTCTGAGAAATTCGACAGGAATATATG	1657
Direct Submission	DB	1059	-----	1058
Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Sukenri-cho, Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp, URL: http://genome.res.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)	QY	1718	TTAAAGAGGACATCTGTAAAGCCCTGTGTGCGTATGATGGAGAAATATGAGACTA	1777
CNA library was prepared and sequenced in Mouse Genome Encyclopedia project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse cDNAs.	DB	1059	-----	1058
URL: http://genome.res.riken.go.jp/	QY	1897	AGTGTGTGAAGCTGTGAAGGCGCCCAAGCCCACTGATGAGGCTGTGCGCGGAGGAC	1837
Direct Submission	DB	1114	AGTGTGTGAAGCTGTGAAGGCGCCCAAGCCCACTGATGAGGCTGTGCGCGGAGGAC	1168
Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Sukenri-cho, Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp, URL: http://genome.res.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)	QY	1938	AGTGTGTGAAGCTGTGAAGGCGCCCAAGCCCACTGATGAGGCTGTGCGCGGAGGAC	1897
CNA library was prepared and sequenced in Mouse Genome Encyclopedia project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse cDNAs.	DB	1169	AGTGTGTGAAGCTGTGAAGGCGCCCAAGCCCACTGATGAGGCTGTGCGCGGAGGAC	1228
URL: http://genome.res.riken.go.jp/	QY	1898	CTTCTTGTGTCCTGTCAGGAGCTGTGCGAGGCGGAGTATCTCATAGAGTGTGCGCTCT	1957
Direct Submission	DB	1229	CGAGTGTGTGAAGCTGTGAAGGCGCCCAAGCCCACTGATGAGGCTGTGCGCGGAGGAC	1288
Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Sukenri-cho, Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp, URL: http://genome.res.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)	QY	1958	GCACACACCGGAGGCTGTGAGGCGGAGTCTGTGAGGCTGTGCGCGGAGGAC	2017
CNA library was prepared and sequenced in Mouse Genome Encyclopedia project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse cDNAs.	DB	1289	GTAGTGTGAAGCTGTGAAGGCGCCCAAGCCCACTGATGAGGCTGTGCGCGGAGGAC	1348
URL: http://genome.res.riken.go.jp/	QY	2018	ASCTGTGACACAGTGTGAAGGCGCCCAAGCCCACTGATGAGGCTGTGCGCGGAGGAC	2077
Direct Submission	DB	1349	ASCTGTGACACAGTGTGAAGGCGCCCAAGCCCACTGATGAGGCTGTGCGCGGAGGAC	1408
Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Sukenri-cho, Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp, URL: http://genome.res.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)	QY	2078	CCGACACACCGGAGGCTGTGAGGCGGAGTCTGTGAGGCTGTGCGCGGAGGAC	2137
CNA library was prepared and sequenced in Mouse Genome Encyclopedia project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse cDNAs.	DB	1409	CAGAGTGTGAAGCTGTGAAGGCGCCCAAGCCCACTGATGAGGCTGTGCGCGGAGGAC	1468
URL: http://genome.res.riken.go.jp/	QY	2138	TAGAGATGTGAAGCTGTGAAGGCGCCCAAGCCCACTGATGAGGCTGTGCGCGGAGGAC	2197
Direct Submission	DB	1469	TTGAGAGGAGGATGTGAAGGCGCCCAAGCCCACTGATGAGGCTGTGCGCGGAGGAC	1528
Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Sukenri-cho, Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp, URL: http://genome.res.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)	QY	2198	CTTGTGCAAAATGTGAAGGCGCCCAAGCCCACTGATGAGGCTGTGCGCGGAGGAC	2257
CNA library was prepared and sequenced in Mouse Genome Encyclopedia project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse cDNAs.	DB	1529	CCATGTGTGCAAGGTTGAAGGCGCCCAAGCCCACTGATGAGGCTGTGCGCGGAGGAC	1588
URL: http://genome.res.riken.go.jp/	QY	2258	TAGATGTGAAGCTGTGAAGGCGCCCAAGCCCACTGATGAGGCTGTGCGCGGAGGAC	2317
Direct Submission	DB	1589	TCGATGGGATTTGCACTGTGGGATGTGACCATGAGCTTGGCTTAAGGCGGTTCTG	1648
Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Sukenri-cho, Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp, URL: http://genome.res.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)	QY	2318	ACGTGTGCGGCTGTGAAGGCGCCCAAGCCCACTGATGAGGCTGTGCGCGGAGGAC	2377
CNA library was prepared and sequenced in Mouse Genome Encyclopedia project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse cDNAs.	DB	1649	ATGATGTGTGTTGTAAGGCGCCCAAGCCCACTGATGAGGCTGTGCGCGGAGGAC	1708

Query Match 15.9%; Score 460; DB 11; Length 4593;
Best Local Similarity 59.9%; Pred. No. 6e-117;
Matches 928; Conservative 0; Mismatches 460; Indels 162; Gaps 3;
QY 1238 TGGGGAAGATGGGACTCGTATGACACCGCATCTTACTGACTGCTCGATATGTT 1297
DB 728 TGGGGAAGATGGGAGAGACAGACACCGCATCTCTCTCAGGATTTGACATTTGTT 787

Qy 2378 CCAAGCACCACACACACACACATATACATGTCACATCTCTTCGGAGCCGGA 2437
 Db 1709 TCACTCAGCAGAAAGCAATGATATACCTGTCTCATCTCCCTCCCTGGGCCCCGA 1768
 Qy 2438 GTATCCGATCATATGAATGAAGCTCTACCTCTCATATCTTGTGGCAATGCCTCA 2497
 Db 1769 CCAATGAGATCAGAGTTCGACCTTTCTTCAGTATCTCTGCTTGAAGCTCAGTC 1828
 Qy 2498 GAGCTACTACCTCAGTACGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 2557
 Db 1829 AAGATTAATCTCAGAGGGCTGAGCATCAGCTGAGCTGAGCTGAGCTGAGCT 1888
 Qy 2558 GCACTACTTCGATCAGAGCTGATATATGAGCTGAGCTGAGCTGAGCTGAGCT 2617
 Db 1889 GACCAAGTTGATATCAGAGCTTCTTACCGCTTGAACACTGTATGACACAGAC 1948
 Qy 2618 CACCAAGCAGATGATGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 2677
 Db 1949 CACAGATGAGAGCTGCTCTTGAATTTGACAGAGAGATCCGGGATCGCTT 2008
 Qy 2678 GGAATATCTCAGCTGCTCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 2734
 Db 2009 GGAATATGACTTCCAGCTGATGAATATCTCAGCTGAGCTGAGCTGAGCT 2068
 Qy 2735 ACATTTGGCCATCTGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 2784
 Db 2069 ACATCTGGCGACAGTGCAGCTCTGACTCTGACTCAGTCACTTGGGTGAGGT 2118

RESULT 4

BM850160
 LOCUS K-EST0130714 S12SN0216 Homo sapiens cDNA clone S12SN0216-81-C06 5',
 mRNA sequence.

ACCESSION BM850160.1 GI:19206559

VERSION EST

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE

AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
 Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
 21C Frontier Korean EST Project 2001

TITLE

Unpublished

JOURNAL

Contact: Kim YS

Genome Research Center

Korea Research Institute of Bioscience & Biotechnology

52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea

Tel: +82-42-860-4470

Fax: +82-42-860-4409

Email: yongseung@mail.kribb.re.kr

Plate: 81 row C column: 06

High quality sequence stop: 625.

FEATURES

SOURCE

Location/Qualifiers

1

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon.9606"

/clone="S12SN0216-81-C06"

/sex="p"

/tissue_type="lymph node"

/cell_type="epithelial"

/cell_line="SNU-216"

/lab_host="top10f"

/clone_lib="S12SN0216"

/note="organ: stomach; Vector: pcns; Site_1: EcoRI;

Site_2: NotI. The poly (A) RNA was dephosphorylated with

bacterial alkaline phosphatase (BAP). Then decapped

with tobacco acid pyrophosphatase (TAP). The decapped

inact mRNA was ligated with DNA-RNA linker including EcoR

I site by treatment of T4 RNA ligase and the first strand
 cDNA was synthesized from oligo dt-selected mRNA by
 priming with dt-tailed vector. The dt-tailed vector was
 circulated to have about 50nt. The cDNA vector was
 EcoRI which site is also included in vector. An RNA strand
 converted to a DNA strand by Okayama-Berg method. The
 obtained cDNA vectors were used for transformation of
 competent cells E. coli Top10f by electroporation method.
 The cDNA libraries constructed by this method are
 full-length enriched cDNA library."

BASE COUNT 181 a 141 c 151 g 152 t

ORIGIN

Query Match 15.8%; Score 456.4; DB 12; Length 625;
 Best Local Similarity 99.8%; Pred. No. 2.7e-116;
 Matches 457; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1230 TGGATTGATGCGGGAAGAATGGAGCTGCTCATGACCCAGCCATCTTACTGACTGGTCTGGA 1289
 Db 168 TGGATTGATGCGGGAAGAATGGAGCTGCTCATGACCCAGCCATCTTACTGACTGGTCTGGA 227
 Qy 1290 TATATGTTCTTGGAGAAATGAGCCCTGTGACACTTTGGGATTTGCCATTAAGTGGAAAT 1349
 Db 228 TATATGTTCTTGGAGAAATGAGCCCTGTGACACTTTGGGATTTGCCATTAAGTGGAAAT 287
 Qy 1350 GTGTAGTAATAATGTCGAGCTGACGATTAATGAGATACAGGCTTTGAGCTTGCCTTCAC 1409
 Db 288 GTGTAGTAATAATGTCGAGCTGACGATTAATGAGATACAGGCTTTGAGCTTGCCTTCAC 347
 Qy 1410 CATTGCCCATGAGTCTGGACACACATTTGGCATGATTCATGATGAGAGAGGAAATGCTG 1469
 Db 348 CATTGCCCATGAGTCTGGACACACATTTGGCATGATTCATGATGAGAGAGGAAATGCTG 407
 Qy 1470 TAAAAGTCTGGAGGGCAACATCATGTCTCCCTCATTTGGCAGAGCCATAGGAGTCTTCTC 1529
 Db 408 TAAAAGTCTGGAGGGCAACATCATGTCTCCCTCATTTGGCAGAGCCATAGGAGTCTTCTC 467
 Qy 1530 CTGGTCACTCTCGAGCCGCGAGTATACACAAATTTCTAAGACCCGCTCAAGCTATCTG 1589
 Db 468 CTGGTCACTCTCGAGCCGCGAGTATACACAAATTTCTAAGACCCGCTCAAGCTATCTG 527
 Qy 1590 CTTTGTGATCAGCCAAAGCTGTGAAGGAATACAGTATCTTGAGAAATTCGCCAGGAGA 1649
 Db 528 CTTTGTGATCAGCCAAAGCTGTGAAGGAATACAGTATCTTGAGAAATTCGCCAGGAGA 587
 Qy 1650 ATTATATGATGCAACACACAGCTGCAAGTGGCAGTTCG 1687
 Db 588 ATTATATGATGCAACACACAGCTGCAAGTGGCAGTTCG 625

RESULT 5

BM845044

LOCUS

DEFINITION

K-EST0123276 S12SN0216 Homo sapiens cDNA clone S12SN0216-97-E10 5',
 mRNA sequence.

ACCESSION

BM845044

VERSION

EST

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE

AUTHORS

Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
 Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
 21C Frontier Korean EST Project 2001

TITLE

Unpublished

JOURNAL

Contact: Kim YS

Genome Research Center

Korea Research Institute of Bioscience & Biotechnology

52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea

Tel: +82-42-860-4470
 Fax: +82-42-860-4409
 Email: yongsung@mail.kribb.re.kr
 Plate: 97 row: E column: 10
 High quality sequence stop: 623.
 Location/Qualifiers
 L: 623

FEATURES

source

/organism="Homo sapiens"
 /mol_type="mRNA 9606"
 /clone_id="S12SN216-97-E10"
 /accession="S12SN216-97-E10"
 /feature_type="Lymph node"
 /cell_type="Epithelial"
 /cell_line="SNU-216"
 /lab_host="Top10F"
 /clone_lib="S12SN216"

/notes="Organ: Stomach; Vector: pCNS; Site 1: EcoRI; Site 2: NotI; The poly (A)+ RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with Labacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including EcoR I site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F, by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

BASE COUNT 181 a 140 c 150 g 152 t
 ORIGIN
 Query Match 15.7%; Score 454.4; DB 12; Length 623;
 Best Local Similarity 39.8%; Pred. No. 9.7e-116;
 Matches 455; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1230 TGGATTGATGGGGAAGATGGGACTCGTCATGACACCCGCTTACTGACTGGTCTGGA 1289
 168 TGGATTGATGGGGAAGATGGGACTCGTCATGACACCCGCTTACTGACTGGTCTGGA 227
 1290 TATATGCTTCTGGGAATGAGCCCTGTGACACTTTGGGATTTGGACCCATTAAGTGGAA 1349
 228 TATATGCTTCTGGGAATGAGCCCTGTGACACTTTGGGATTTGGACCCATTAAGTGGAA 287
 1350 GTGTAGTAATATGTCAGCTCCGATGATATGAGATACAGTCTTGGACTGGCTTCAC 1409
 288 GTGTAGTAATATGTCAGCTCCGATGATATGAGATACAGTCTTGGACTGGCTTCAC 347
 1410 CATTCCTCCATGAGCTTCGACACATCTTGGCTGATGATGATGAGGAGGAGGACATG 1469
 348 CATTCCTCCATGAGCTTCGACACATCTTGGCTGATGATGATGAGGAGGAGGACATG 407
 1470 TAAAGAGTCGAGGACACATCATGTCCTCATGATGGCAGGACGATGATCTTC 1529
 408 CAAAGTCTCGAGGACACATCATGTCCTCATGATGGCAGGACGATGATCTTC 467
 1530 CTGCTCAGCTCAGGCGCCGATGATGATGATGATGATGATGATGATGATGATGATG 1589
 468 CTGCTCAGCTCAGGCGCCGATGATGATGATGATGATGATGATGATGATGATGATG 527
 1590 CTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1649
 528 CTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 587
 1650 ATTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1685
 588 ATTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 623

RESULT 6

BM808410

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BM808410 1092 bp mRNA linear EST 05-MAR-2002
 AGNCOURT_6604069 NIH_MGC_125 Homo sapiens cDNA clone IMAGE:5591626
 S', mRNA sequence.

BM808410.1 GI:19125233

EST.

Homo sapiens (human)

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

BM808410

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BM808410 1092 bp mRNA linear EST 05-MAR-2002
 AGNCOURT_6604069 NIH_MGC_125 Homo sapiens cDNA clone IMAGE:5591626
 S', mRNA sequence.

BM808410.1 GI:19125233

EST.

Homo sapiens (human)

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

BM808410

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BM808410 1092 bp mRNA linear EST 05-MAR-2002
 AGNCOURT_6604069 NIH_MGC_125 Homo sapiens cDNA clone IMAGE:5591626
 S', mRNA sequence.

BM808410.1 GI:19125233

EST.

Homo sapiens (human)

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

BM808410

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BM808410 1092 bp mRNA linear EST 05-MAR-2002
 AGNCOURT_6604069 NIH_MGC_125 Homo sapiens cDNA clone IMAGE:5591626
 S', mRNA sequence.

BM808410.1 GI:19125233

EST.

Homo sapiens (human)

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens


```

Db      319  GAGGAAACCGGGTGTGGCTGGGAAATCTCATGCTCGCTTGGGACCGAGAGCAGC 378
Qy      2717  CCCCTGCCAGCCAGCCAGCTACCTTGGGCCATCGTGGCTCTGAGCTCTCCCTGTCCTGG 2776
Db      379  CCCCTGCCAGCCAGCCAGCTACCTTGGGCCATCGTGGCTCTGAGCTCTCCCTGTCCTGG 438
Qy      2777  GAGGGGG 2783
Db      439  GAGGGGG 445

RESULT 7
BX458327
LOCUS   BX458327 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE001YN22
DEFINITION 5-PRIME, mRNA sequence.
ACCESSION BX458327.1 GI:31028976
VERSION   BX458327.1
KEYWORDS EST.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE    1 (bases 1 to 962)
COMMENT  Li, W.B., Gruber, C., Jesse, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqreg@genoscope.cns.fr, web : www.genoscope.cns.fr
Library was constructed by life technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 8035.r For
more information, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DE001DCL1Q1P1&cluster=8035.r. Contact :
B. Liang, Email: b.liang@life.technet.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DE001DCL1Q1P1.
FEATURES
Source Location/Qualifiers
1..962
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE001YN22"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."
BASE COUNT 268 a 202 g 232 t
ORIGIN
Query Match 13.9%; Score 401.2; DB 13; Length 962;
Best Local Similarity 65.9%; Pred. No. 9.7e-101;
Matches 588; Conservative 5; Mismatches 298; Indels 1; Gaps 1;
Qy 1542 CAGCGCCAGTATCTACACAAATTCATAGCACCGCTCAAGCTATCTGCTTGTCTGATCA 1601
Db 71 CAGCGCCAGTATCTACAGAAATTCCTAGCACACCTCAAGCGGGGCTGTAGTGATGA 130
Qy 1602 GCGAAGCGCTGTAGAGGATACACATATCTCTGAAATTCGACGAGATTTATATGATGC 1661
Db 131 GCCCAGCAGACGACGACGATATATATTCGCGACCACTACGACGAGATTTATGTC 190
Qy 1662 AACACACAGTGCAGTGCAGTCTCGAGAGAACCCAGCTCTGCTGCTGCTGCTTAA 1721
Db 191 TGCACACAGTGAATGGCAATTTGGAGCAACCCAGTATGACGCTTGTGTGT 250
Qy 1722 AACGACATCTGAAGCCTGTGCTCCATCGTATTGGAAGGAATGTGAGACTAATTT 1781

```

```

Db      251  GAAGGATATTTGCAATCACTATTGTGTGCCACCGAGTAGGCCACAGAGTGTGAGACCAAGTT 310
Qy      1782  TATGCCAGCAGCAGAGGACACAAATTTGTGGGATGATCATGTGTGGTGGGAGGACAGTGTG 1841
Db      311  TATGCCAGCAGCAGAGGACACAAATTTGTGGGATGATCATGTGTGGTGGGAGGACAGTGTG 370
Qy      1842  TGTCAATATATGATGATGAGAGGCCCCCAAGCCCAATGATGAGCCCACTGGTGGGCTGCTTC 1901
Db      371  CATAAAGTTTGGGAGCTCGGGGCCCGGCCCATCCAGGCCAGTGGTGGTGGGCTGGGACAA 430
Qy      1902  TTGGTCCCATCTCTCCAGGACCTCGGGAGGGGAGTATCTCATAGGAGTCCGCTCTGACAC 1961
Db      431  GTGGTCAGAAATGTTCCCGGACATGTGGTGGAGAGTCAAGTTCGAGAGACACACTGCNA 490
Qy      1962  CAACCCCAAGCCATCGATGAGGAGGAAATTTCTGTGAGGGCTCCACTCGCACTCTGAAGCT 2021
Db      491  TAAACCCCAARCTCAGTATGTGGCATATTTCTGTCAGAGTTCCTAGCCGCTATTATCAGCT 550
Qy      2022  CTGCAACAGCTCAGAAATGTCCCGGACAGTGTGTGACTTCCGCTGCTGCTCATGTGGCCA 2081
Db      551  GTGCATATATTACCCCTTCGATGAAATAGCTTGGATTTTCGGGCTCAACAGTGTGGCA 610
Qy      2082  GCACACAGCAGACGATTCAGAGGCGGCGACATACAGTGGAGAGCTTTCACCTCAAGTAGA 2141
Db      611  ATATACACGACAAACCTTTCGGTGGATGGTTCTTACCACTGGAAACCTTATCGAAAGTGA 670
Qy      2142  AGTCAGGACTTATGCAACTCTACTGTATGCAAGAGGATTTG-ATTTCCTCTTCTT 2200
Db      671  AGGAGAGATTCGATCAAGAGTGTACTGAGCTGTGAGACTTGAATTTGAAATTTTTCGAA 730
Qy      2201  TGTCAATATGCTGAGAGTGGATCTCGATGCTCGAGAGTACCGGTATGTTGATAG 2260
Db      731  TGTCCGCAATGTGAAAGGAGATCTCCCTGCTCCCAACAAAGATGTTTATTTG 790
Qy      2261  ATGGGATGATGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2320
Db      791  ACAGGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 850
Qy      2321  TCTGTGGGTTGTAAACGGATTAACCTACAGCTCAGATTCACAGGGGCTCTTACACCA 2380
Db      851  CTGTGGGTTTGCAGAGTGAATTAATCACTTCGAAATTTTATAAAGGCTGTACTCTCA 910
Qy      2381  AGCACACACACACACACACATTAATCAAGTGGTCAACATTCCTTCTGGAGC 2432
Db      911  ACCAGCATTAACCAAGATATATATTCGGTGTGTCATCATCTCCAGCTGGCGC 962

RESULT 8
BX45406
LOCUS   BX45406
DEFINITION K-EST0123712 S12SN216 Homo sapiens cDNA clone S12SN216-99-B04 5',
mRNA sequence.
ACCESSION BX45406
VERSION   BX45406.1 GI:19201805
KEYWORDS EST.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE    1 (bases 1 to 537)
COMMENT  Kim, N.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R.,
Oh, K.J., Cheong, J.E., Sohn, H.Y., Kim, J.N., Park, H.S., Kim, S. and
Kim, Y.S.
21C Frontier Korean EST Project 2001
Unpublished
Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Soeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Page: 99 row: 8 column: 09

```

High quality sequence stop: 537.
 FEATURES
 source
 Location/Qualifiers
 1. 537
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="S12SNU216-99-804"
 /sex="F"
 /tissue_type="Lymph node"
 /cell_type="Epithelial"
 /cell_line="SHU-216"
 /lab_host="Top10F"
 /clone_lib="S12SNU216"
 /notes="Organ: Spleen; Vector: pCNS; site 1: EcoRI; site 2: NotI; The poly (A) + RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tobacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F, by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."
 BASE COUNT 149 a 124 c 129 g 135 t
 ORIGIN
 Query Match 12.7%; Score 368.4; DB 12; Length 537;
 Best Local Similarity 99.7%; Pred. No. 1.2e-91;
 Matches 369; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1230 TGGATGATGGGGAAGATGGGCTCATGACACCGCCATCTTACTGCTGCTGGA 1289
 DB 168 TGGATGATGGGGAAGATGGGCTCATGACACCGCCATCTTACTGCTGCTGGA 227
 QY 1290 TATATGCTCTGGGAGATGCCCTGTGACATTTGGGATTTGCACCATTAAGTGGAT 1349
 DB 228 TATATGCTCTGGGAGATGCCCTGTGACATTTGGGATTTGCACCATTAAGTGGAT 287
 QY 1350 GTGTAGTAAATATCGCAGCTCGACATTAATGAAGATACAGCTTTGGAGCTTCCTC 1409
 DB 288 GTGTAGTAAATATCGCAGCTCGACATTAATGAAGATACAGCTTTGGAGCTTCCTC 347
 QY 1410 CATTGCCCATGCTTGGACACACTTTGGCATGATTCATGATGAGAGGGAACATG 1469
 DB 348 CATTGCCCATGCTTGGACACACTTTGGCATGATTCATGATGAGAGGGAACATG 407
 QY 1470 TAAAGATCCGAGGCAACATCATGTCCTTACATTTGGCAGGCGCAATGATGCTTC 1529
 DB 408 CAAAGATCCGAGGCAACATCATGTCCTTACATTTGGCAGGCGCAATGATGCTTC 467
 QY 1530 CTGTCACCTCGAGCGCGCAGTATCTACGAAATTTTCAGACCGCTTCAAGCTATCTG 1589
 DB 468 CTGTCACCTCGAGCGCGCAGTATCTACGAAATTTTCAGACCGCTTCAAGCTATCTG 527
 QY 1590 CTTTGTGAT 1599
 DB 528 CTTTGTGAT 537
 RESULT 9
 BQ44190
 LOCUS
 DEFINITION BQ44190 698 bp mRNA linear EST 29-MAY-2002
 IMAGE:5709081 5', mRNA sequence.
 ACCESSION BQ44190
 VERSION BQ44190.1 GI:21547302
 KEYWORDS EST.

Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NIH-MGC http://mgs.nci.nih.gov/
 1 (bases 1 to 698)
 TITLE
 JOURNAL
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: gcpbs-@mail.nih.gov
 Tissue Procurement: Dr. James Lin, University of Iowa
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
 This clone was contributed by the Brain Molecular Anatomy Project (BNAP)
 Seq primer: pYX-5.
 Location/Qualifiers
 1. 698
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:5709081"
 /tissue_type="whole brain"
 /dev_stage="embryo 15.5 dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH_BMAP_EXO"
 /notes="Organ: brain; Vector: pYX-Asc; Site 1: EcoRI; Site 2: Not I; the library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with EcoRI adaptor, digested with Not I, and then ligated directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail, is GTGGTGGA. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BNAP). Gene Discovery in the Developing Mouse Nervous System, supported by National Institutes of Mental Health (NIMH), Remington, Ph.D., program coordinator."
 BASE COUNT 188 a 166 c 189 g 155 t
 ORIGIN
 Query Match 11.7%; Score 339.2; DB 13; Length 698;
 Best Local Similarity 72.3%; Pred. No. 2.1e-83;
 Matches 434; Conservative 0; Mismatches 158; Indels 0; Gaps 0;
 QY 1238 TGGGGAAGTGGGAGCTGTCATGACACCGCCATCTTACTGCTGCTGATATATGTT 1297
 DB 107 TGGGGAAGTGGGAGCTGTCATGACACCGCCATCTTACTGCTGCTGATATATGTT 166
 QY 1298 CTTGGAAGATGAGCCTGTGACATTTGGGATTTGCACCCATAAGTGAATGTAGTA 1357
 DB 167 CTTGGAAGATGAGCCTGTGACATTTGGGATTTGCACCCATAAGTGAATGTAGTA 226
 QY 1358 AATATCGCAGCTGCAGCATTAATGAAGATACAGCTTGGAGCTGGCTTACCATTTGCC 1417
 DB 227 AATATCGCAGCTGCAGCATTAATGAAGATACAGCTTGGAGCTGGCTTACCATTTGCC 286
 QY 1418 ATGAGTCTGGACACAACTTTGGCATGATTCATGATGGAGAGGAACATGTGTAAAGT 1477
 DB 287 ATGAGTCTGGACACAACTTTGGCATGATTCATGATGGAGAGGAACATGTGTAAAGT 346
 QY 1478 CCGAGGCAACATCATGTCCTTACATTTGGCAGACGCAATGTGAGTCTTCTCTGTCAC 1537
 DB 347 CAGAGGCAACATCATGTCCTTACATTTGGCAGACGCAATGTGAGTCTTCTCTGTCAC 406

QY 1538 CCTGAGCGCCAGTATCTACAAATTTCTAAGCAGCGCTCAAGCTATCTGCTGTG 1597
 Db 407 CCTGAGCGCCAGTATCTAAGAAATTTCTCAGTACACCGCAGGCTGCTGTG 466
 QY 1598 ATCAGCCAAAGCCTGTGAAGGAATCAAGTATCTGAGAAATTTGCGAGAGATTAATG 1657
 Db 467 ATGAGCCCAAGCAACCGGACAGTATAAATATCCGACAACTCCCGGACAGATTTATG 526
 QY 1658 ATGCAAAACACACAGTGCAGTGCAGTTCGGAGGAAAGCCAGCTCTGCACTCTGGACT 1717
 Db 527 ATGCGACATGCAATGTGAAGTGCAGTATTTGGACCAAGCAAGCTGTGAGGCTTGGG 586
 QY 1718 TTTAAAGGACATCTGTAAGCCCTGTGTCCTGTCATCTGTAAGGAAATGTGAGACTA 1777
 Db 587 TTTAAGGACATCTGTAAGTGCAGTATTTGGACCAAGCAAGCTGTGAGGCTTGGG 646
 QY 1778 AATTATGTCAGCAGCAGAGGACCAATTTTGGGCAATGATGTTGGGCG 1829
 Db 647 AGTTCATGCTGTCAGCAGAGGAGCGCTTGTGCTTGAATGTGTTGTCG 698

RESULT 10

BM906555 1122 bp mRNA linear EST 12-MAR-2002
 AGENCOURT_6621721 NIH_MGC_125 Homo sapiens cDNA clone IMAGE:5722286
 5', mRNA sequence.

ACCESSION

BM906555

VERSION

BM906555.1 GI:19356934

KEYWORDS

Homo sapiens (human)

SOURCE

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

NIH-MGC <http://mgs.mci.nih.gov/>

AUTHORS

Wang, S. et al. 1122

TITLE

NIH-MGC <http://mgs.mci.nih.gov/>

COMMENT

Cloned by Robert Strausberg, Ph.D.
 Tissue procurement: Invitrogen
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LHAM12708 row: b column: 15
 High quality sequence start: 19
 High quality sequence stop: 633.

FEATURES

source

1..1122
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5722286"
 /lab_host="DH108"
 /clone_lib="NIH MGC 125"
 /notes="Organ: ovary (pool of 3); Vector: pCMV-SPORT6;
 Site 1: EcoRV (destroyed); Site 2: NotI; RNA source pool
 of three ovaries, from females ranging in age from 38 to
 49 yo. Library is oligo-dT primed and directionally cloned
 (EcoRV site is destroyed upon cloning). Average insert
 size 2.1 kb, insert size range 1-3.5 kb. Library is
 normalized and enriched for full-length clones and was
 constructed by C. Gruber (Invitrogen). Research Genetics
 tracking code 036." 233 t
 BASE COUNT 296 a 306 c 287 g 233 t

ORIGIN

Query Match 11.7%; Score 338.2; DB 12; Length 1122;
 Best Local Similarity 70.6%; Pred. No. 4.8e-83;
 Matches 581; Conservative 0; Mismatches 198; Indels 44; Gaps 8;

QY 428 GCGATTACGTGTCCCATGAATCATGCCCATCAGCGCGGAGAGAGCAGTGGCGGTGT 487
 Db 1 GCGATTACGTGTCCCATGAATCATGCCCATCAGCGCGGAGAGAGCAGTGGCGGTGT 60
 QY 488 CCGAGGTGTGAGTCTCTTCACTCTCGCTGGAAGGCCCGCCAGGACAGCTTCCACATGGATC 547
 Db 61 CCGAGGTGTGAGTCTCTTCACTCTCGCTGGAAGGCCCGCCAGGACAGCTTCCACATGGATC 120
 QY 548 TGAGGACTTTCAGCAGCCCTAGTGGCTCTCTGGCTTTTATTGTCCAGAGCTTG--GGAAGACA 606
 Db 121 TCGGAGCTTCAGCAGCCCTAGTGGCTCTCTGGCTTTTATTGTCCAGAGCTTGCGGAAGACA 180
 QY 607 GGCACTAAGTCTGTGCAGACTTTACCGCCAGAGGACTTCTGTTTCTATCAAGGCTTTTG 666
 Db 181 GGCACTAAGTCTGTGCAGACTTTACCGCCAGAGGACTTCTGTTTCTATCAAGGCTTTTG 240
 QY 667 GATTCACACAGAAATCGCCATCGCATGGAAGGAACTTCTGTGAGGCTTCCACTCGCACT 726
 Db 241 GATTCACACAGAAATCGCTCA-----GTGGCCCTTTCAAGC 276
 QY 727 CTGAAGCTCTGCAACAGTCAAGATGTCCCGGACAGTGTGTGACTTCTGCTGTGCTGAG 786
 Db 277 TGCCCAAGGCTTGTCCAGGCATGATGAAACAGAGAGGCGCAGATTAATTCTTAAGGCCACTT 336
 QY 787 TGTCGAGCACAACAGCGAGAGGATTCAGAGGCGGCGCATCAAGTGG--AAGCCTTACA 844
 Db 337 CTTCAACACCTCTCTGGAAGACTCGGACAGCTGCCCAAGGACGTGCGCATCCACAGTA 396
 QY 845 CTCAGTAGAAGCGGCTTATGCAACTCTGATCTGATATCGAGAGAGGATTAATTTCTTCT 904
 Db 397 CTGTCAGAGATTCACAGGCCCTGCTCTGGGCGCAGTGAAGTCTGTGACCTC- 455
 QY 905 TTCTTTGTCATTAAGTCAAGATGSGACTCCTGCTGCGAGGATAGCTGTAATGTTT 964
 Db 456 -----AAGACATGGGCTGSCATCACTCCCTGACAGAGGACCTTGC 504
 QY 965 GTATGATGGATGTGTGAGCTCAGTGTGGTGTCCCATCTG-CGCACTGCCCGACCT 1023
 Db 504 CTGGACTGCCAAGACACATTTCTGTGAAGAGCGCAAGATACATGCCACACT 564
 QY 1024 CCGAGAGAGACCTCTCTATCTCCAGATGATTAAGTATGCTTCTGCTGAGCAGTACGCG 1083
 Db 565 CCGAGAGAGACCTCTCTATCTCCAGATGATTAAGTATGCTTCTGCTGAGCAGTACGCG 624
 QY 1084 TCTCTCTGAGGTCCCATAGAAATGAAGAACTGAAGAACTGAAGTGGAGACCTTGTGGGCTCGAC 1143
 Db 625 TCTCTCTGAGGTCCCATAGAAATGAAGAACTGAAGAACTGAAGTGGAGACCTTGTGGGCTCGAC 684
 QY 1144 AAAAAGATGATGCAAAA--CCATGCGCATGAAAATATCA--CCACCTAGTGTCTACCATAC 1201
 Db 685 AAAAAGATGATGCAAAAACCCATGCGCATGAAAATATCAACCCCTTACGGGCTGACCATAC 744
 QY 1202 TCACATGCTGA--TCGTCTTTTATCAAGATGATTAAGTGGG 1242
 Db 745 TCACATGCTGAATCTGCTTTTATCAAGAGGAGCAATAGG 787

RESULT 11

BF933693/c

LOCUS

IL5-NT0228-211200-358-e08 NT0228 Homo sapiens cDNA, mRNA sequence.

DEFINITION

BF933693

ACCESSION

BF933693

VERSION

BF933693.1 GI:12351017

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 332)

DIAS Neco.R., Garcia Correa.R., Verjovski-Almeida.S., Briones.M.R.,

Nagai.M.A., da Silva.W. Jr., Zago.M.A., Bordin.S., Costa.F.,

Goldman.G.H., Carvalho.A.F., Matsukuma,A., Baia.G.S., Simpson,D.H.,

Brunstein.A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare

QY 1650 ATTATGATGCAACACACAGTGCAGTGCAGTTCGGAGAGAGCCAGCTGCTCAT 1709
 DB 251 GATTATGATGCGACACGACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 310
 QY 1710 GCTGGACTTTTAAAGGACATCTGTAAAGCCCTGTGTGTGTGTGTGTGTGTGTGT 1769
 DB 311 CTTTGGGGTATGAGAGCAATTTGCAAGTGTCTGTGTGTGTGTGTGTGTGTGTGT 370
 QY 1770 TGAGATCAAAATTTATCCAGCAGCAGAGAGGCAAAATTTGTGGGATGATGTGTGTGTGT 1829
 DB 371 TGAAACCAAGTTTCATCCCTGCAGCCGAGGCGACCGTCTGTGTGTGTGTGTGTGTGT 430
 QY 1830 GAGAGAGACAGTGTGTGAATATGTGTGAAGGCCCAAGCCCAAGCCCAAGCCCAAGTGT 1889
 DB 431 GAAAGCCCAATGT 490
 QY 1890 GGAATGT 1949
 DB 491 TGCTGT 550
 QY 1950 TGCCTCTGTGCACCAACCCCAAGCCATGCGATGAGAGGAGTGTGTGTGTGTGTGTGT 2009
 DB 551 GAGACACTGT 610
 QY 2010 CACTGTGAGT 2069
 DB 611 CATCTACAAGT 670
 QY 2070 TGT 2129
 DB 671 GCAATGT 730
 QY 2130 CA 2131
 DB 731 TA 732

RESULT 13
 B0844919 430 bp mRNA linear EST 06-MAR-2002
 LOCUS K0870123130 S125NU216 Homo sapiens cDNA clone S125NU216-91-A08 5',
 DEFINITION K0870123130 S125NU216 Homo sapiens cDNA clone S125NU216-91-A08 5',
 ACCESSION B0844919
 VERSION B0844919
 KEYWORDS EST
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 430)
 Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
 Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
 Kim,Y.S.
 21C Frontiers Korean EST Project 2001
 Unpublished
 Contact: Kim YS
 Genome Research Center
 Korea Research Institute of Bioscience & Biotechnology
 52 Eosun-dong Yusong-gu, Daejeon 305-333, South Korea
 Tel: +82-42-860-4470
 Fax: +82-42-860-4409
 Email: yongsungemail.kribb.re.kr
 Plate: 91 row: A column: 08
 High quality sequence stop: 430.
 Location/Qualifiers
 1. .430
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="S125NU216-91-A08"
 /sex="F"
 /tissue_type="Lymph node"
 /cell_type="Epithelial"

FEATURES

source

/cell_line="SNU-216"
 /lab_host="Top10f"
 /clone_lib="S125NU216"
 /notes="Organ: Stomach; Vector: pCNS; site 1: EcoRI;
 site 2: NotI; The poly (A)+ RNA was dephosphorylated with
 bacterial alkaline phosphatase (BAP) and then decapped
 with Cabaco acid pyrophosphatase (TAP). The decapped
 intact mRNA was ligated with DNA-RNA linker including EcoR
 I site by treatment of T4 RNA ligase and the first strand
 cDNA was synthesized from oligo dT-selected mRNA by
 priming with dT-tailed vector. The dT-tailed vector was
 adjusted to have about 60nt. The cDNA vector was
 circularized with E. coli DNA ligase after digestion of
 EcoRI which site is also included in vector. An RNA strand
 converted to a DNA strand by Okayama-Berg method. The
 obtained cDNA vectors were used for transformation of
 competent cells E. coli Top10f, by electroporation method.
 The cDNA libraries constructed by this method are
 full-length enriched cDNA library."
 BASE COUNT 125 a 89 c 108 g 108 t
 ORIGIN

Query Match 9.0%; Score 259.8; DB 12; Length 430;
 Best Local Similarity 99.2%; Pred. No. 3.2e-61;
 Matches 261; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1230 TGGATGATGAGGAGATGGGACTGTCTATGACACACCGCCCTTACTGACTGCTTGA 1289
 DB 168 TGGATGATGAGGAGATGGGACTGTCTATGACACACCGCCCTTACTGACTGCTTGA 227
 QY 1290 TATATGCTTGGAGATGAGCCCTGTGACATTTGGGATTTGCCCTTACTGAGT 1349
 DB 228 TATATGCTTGGAGATGAGCCCTGTGACATTTGGGATTTGCCCTTACTGAGT 287
 QY 1350 GTCTGTGATGATGAGGATGAGGATTTGAGGATGAGGATGAGGATGAGGATGAGGAT 1409
 DB 288 GTGTGATGATGAGGATGAGGATTTGAGGATGAGGATGAGGATGAGGATGAGGAT 347
 QY 1410 CATTTGCCATGCTGACGACACATTTGGGATTTGAGGATGAGGATGAGGATGAGGAT 1469
 DB 348 CATTTGCCATGCTGACGACACATTTGGGATTTGAGGATGAGGATGAGGATGAGGAT 407
 QY 1470 TAAAGATGCCAGGCGGACACATCA 1492
 DB 408 CTAAAGATGCCAGGCGGACACATCA 430

RESULT 14
 B0844919 657 bp mRNA linear EST 11-NOV-2002
 LOCUS K0870123130 S125NU216 Homo sapiens cDNA clone S125NU216-91-A08 5',
 DEFINITION K0870123130 S125NU216 Homo sapiens cDNA clone S125NU216-91-A08 5',
 ACCESSION B0844919
 VERSION B0844919
 KEYWORDS EST
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 657)
 Satou,Y., Shin-I,T., Kohara,Y. and Satoh,N.
 Expressed genes in Ciona intestinalis (2002c)
 Unpublished
 Contact: Nori Satoh
 Department of Zoology
 Kyoto University
 Sakyo-ku, Kyoto 606-8502, Japan
 Tel: 81-75-753-4081
 Fax: 81-75-705-1113
 Email: satoh@ascidian.zool.kyoto-u.ac.jp.
 Location/Qualifiers
 1. .657
 /organism="Ciona intestinalis"

Kanagawa 230-0045, Japan (E-mail: genome-res@qsc.riken.go.jp, genome-res@qsc.riken.go.jp, Tel: 81-45-503-9222, Fax: 81-45-503-9216).

Accession Number: U00001
 Location/Qualifiers
 1..3399
 organism="Mus musculus"
 mol_type="mRNA"
 strain="C57BL/6J"
 db_xref="FANTOM,DB:9630005B22"
 db_xref="taxon:10090"
 clone="9630005B22"
 tissue_type="cerebellum"
 clone_library="Riken full-length enriched mouse cDNA library"
 dev_stage="16 days neonate"
 364..1737

264. 1757
 notes=unmated protein product; hypothetical
 metallothionein (M128) family
 containing protein [InterPro:IPR001590, evidence:
 InterPro]
 putative"
 codon_start=1
 protein_id="G6C29190.1"
 db_xref="GI:26330920"
 translation="MEALCICLCALRAAGPPGPGGLGRLAKLQICCPCCASVAV
 LALASGSGSGGLNDIVFVWFGSGYSITSDILHHRKRSAGASNSLRYRS
 AGQDHLBELFSLSHFRVPLQKDGASERFVFPCLQYGFINDSSSVAS
 CAGLGLKIDTDFEFLISLPQLLGHNSPPAGHPVLLYKTAERKRYNQDY
 ASQSTYTGASGHPSTPAQSQEPYSHRKRKQCKRKKYKLPAPDAYLRDPEY
 QGQGRFARGASQNGVNLVADAMVRKGDYTVYLLVANNVSLSEKPEY
 SDNSINIVVSLLEPEGLLNHSDQSLNSFCOMQSLVSGKSGKHDAILLTGF
 IISLQKSNKNEPQDLPFAISQMSKSTINEDTGLGLATFAHESGNGFQMHDEG
 INCRKAGHINSPLTNGNGVSWSSCSQRYLKKFLRTFCIAFTG"

BASE COUNT	905 a	838 c	774 g	882 t	ORIGIN
Query Match	7.7%	Score 221.6;	DB 11;	Length 13399;	
Best Local Similarity	76.4%;	Prod No. 3.9e-50;			
Matches 27;	Conservative	0;	Mismatches 84;	Indels 0;	Gaps 0;
1238 TGGGGAGATGGGACTGGTCATGACGACCGACCATCTTATGACTGCTCTGATATATGTT	1297				
1390 TGGGAAGATGGCAGAGAGACACGACCGGCCATCCTCTCTCACGAGATTGACATTTGTT	1449				
1298 CTTGGAAGATGAGCCCTGGTGACACTTTTGGGATTTTGACCCATTAAGTGGAGATGTTAGTA	1357				
1450 CCTGGAAGACGAACCATGTGACACACTAGGATTTGCTCTATCAGTGGCATGTGCAGTA	1509				
1358 AATATGCGAGCTGCAGATTAATGAAGATACAGGTCTTGGACTGGCCTTCACACTTGGCC	1417				
1510 AGTACGGAAGCTGTACCATCAATATGAGACACAGGACTTGGCCTTGGCCTTCACACTTGGCC	1569				
1418 ATGAGTCTGGACACAACTTTGGCATGATTCATATGATGGAGAGGAAACATGTGTAAAGAT	1477				
1570 ATGAGTTCAGGSCAACCTTTGGCATGGTTTCATATGTTGGTGAAGCAATCCCTGCGAGGAAG	1629				
1478 CCGAGGGGCACATCATGTCCCTCATATTGGCAGGAGCGCATTTGGAGTGTCTTCCTGGTCAC	1537				
1630 CAGAAGGCCACATCATGTCAACCCNACACTGACTGGAAACATNTGGGTTGTTTTCATGCTGT	1689				
1598 CCTGAGCGCGCAGTATGTACACAAATTTCTTAGCACCGCTCAAGCTATCTGCGTT	1593				
1690 CCTGAGCGCGCAGTATGTAAAGAAATTTCTTCAGAACTCTTAGATCTATACATTT	1745				

Search completed: October 28, 2003, 23:23:11

OM nucleic - nucleic search, using sw model
Run on: October 28, 2003, 18:26:52 ; Search time 194 Seconds
(without alignments)
6586.619 Million call updates/sec

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Total number of hits satisfying chosen parameters: 1139956

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45%

```
1: /cgn2_6/pdata/2/ina/5A_COMB.seq.*
2: /cgn2_6/pdata/2/ina/5B_COMB.seq.*
3: /cgn2_6/pdata/2/ina/6A_COMB.seq.*
4: /cgn2_6/pdata/2/ina/6B_COMB.seq.*
5: /cgn2_6/pdata/2/ina/pCTUS_COMB.seq.*
6: /cgn2_6/pdata/2/ina/backfiles.seq.*
```

SUMMARIES

Result No.	Score	Query Match	DB	ID	Description
C	1	1874.6	64.8	4	US-09-930-972-2
	2	1874.6	64.8	4	US-09-930-972-5
	3	1874.6	64.8	4	US-09-930-972-1
	4	1874.6	64.8	4	US-09-930-972-1
	5	1824.6	61.3	3218	US-09-369-364-5
	6	176.8	6.1	3885	US-09-369-364-6
	7	176.8	6.1	3885	US-09-369-364-16
	8	118.4	4.1	2502	US-09-369-364-4
	9	118.4	4.1	2502	US-09-491-522-2
	10	114.2	3.9	2450	US-09-491-522-1
C	11	114.2	3.9	2450	US-09-491-522-9
	12	113.6	3.9	5804	US-09-491-522-8
	13	102.4	3.5	3002	US-09-369-364-12
	14	101.2	3.5	2184	US-09-369-364-1
	15	97.2	3.4	3706	US-09-445-0233-13
	16	97.2	3.4	4756	US-09-489-9708-58
	17	97.2	3.4	4858	US-09-130-491-1
	18	94	3.2	2184	US-09-395-184-1
	19	85.2	2.9	2625	US-09-445-0233-2
	20	85.2	2.9	3358	US-09-369-364-14
C	21	83.6	2.9	3350	US-09-369-364-8
	22	83.6	2.9	3350	US-09-125-1268-14
	23	77.2	2.6	29	US-09-634-2866-14
	24	76.4	2.6	25464	US-09-326-4800-4
	25	76.2	2.6	32042	US-09-326-4800-4
	26	76.2	2.6	32042	US-09-245-281-44
	27	75.6	2.6	32042	US-09-245-281-44
	28	75.6	2.6	1767	US-09-620-6203-63
	29	75.6	2.6	1767	US-09-620-3120-247
	30	75.6	2.6	1767	US-09-620-3120-247

Db 349 AGCCATAGTGGCTTCTTCTTATTTGTCAGAGCGTGGGAAGAAGAGCCATAGTCTGTG 408
 Qy 622 CAGACTTTCACCGCCAGAGGACCTCTGTTTCATCAGAGGCTCTTTGGGATCAGCAGAAAC 681
 Db 409 CAGACTTTCACCGCCAGAGGACCTCTGTTTCATCAGAGGCTCTTTGGGATCAGCAGAAAC 468
 Qy 682 TGGCATCGGATCGAGGGAAGTCTGTGAGGCTCTCACTCGGACTCTGAGCTCTGCAC 741
 Db 469 TCTCTCA -----GTGGCCCTTTCACCTCGCCAGGCTGTGCA 804
 Qy 742 AGTCAGAAATGTCCTCGGAGCAGAGTGTGACTTCGGTCTCGTCTGAGTGTGCGGAGCAGCA 801
 Db 505 GGCATGATACGAACAGAGAGGCGAGATTAATCTTCAAGGCCACTTCTTCACACTCTCA 564
 Qy 802 AGCAGACGATTCAGAGGGCGGCACTCAAGTGG - AAGCTTACACTCAAGTAGAGAGCG 859
 Db 565 TGGAACTCGGCAAGCTGCCAGGCGAGCTCGCCATCCACGACTGTGTACAGAGATCC 624
 Qy 860 ACTTATGCAAACTCTACTGTATTCGAGAGGATTTGATTTCTTCTTTGTGCAATA 919
 Db 625 ACAGAGCCCATGCTCTCGGCGCAGTGAAGTCTGTGAGCTC -----AAGG 672
 Qy 920 AAGTCAAGATGGGACCTCATGCTCGAGGATAGCTGTAATGTTGTATAGATGGATAT 979
 Db 673 ACATGGAGCTGGCACAATCAACCTCTGCACAGAGCGACTTCGCCCTGGGACTGCCA 732
 Qy 980 GTGAGCTCAGTGTGTGTCACATCTGC - GCACATGCCCGAGCTCCAGGAGAGACCTC 1038
 Db 733 AAGCAGCATTTCTGTGGAAGAGCGCAAGAAATACATGCCCGAGCTCCAGGAGAGACCTC 792
 Qy 1039 TTCACTTCGCGAGATGATAGTCTGCTTACGGCATTAAGCGCTCTCTCTGAGGTC 1098
 Db 793 TTCACTTCGCGAGATGATAGTCTGCTTACGGCATTAAGCGCTCTCTCTGAGGTC 852
 Qy 1099 CATAGAAATGAAGACATGAACCTGGAGACTTGTGTGTGTCGACAAAAGATGATGCA 1158
 Db 853 CATAGAAATGAAGACATGAACCTGGAGACTTGTGTGTGTCGACAAAAGATGATGCA 912
 Qy 1159 AACATGGCCGATGAATATGACACTGATGCTCACATCTCACATGATGATGCT 1218
 Db 913 AACATGGCCGATGAATATGACACTGATGCTCACATCTCACATGATGATGCT 972
 Qy 1219 TTATTCAAAG ----- 1229
 Db 973 TTATTCAAAGATGAACATAGAGGAACATCAACATTCGAAATTTAGTCTGATCTT 1032
 Qy 1230 ----- 1229
 Db 1033 CTAGAGATGAACAGCCAGGACTGGTGTATAGTCAACAGCAGACACACTTAAGTAGC 1092
 Qy 1230 -----TGGATTAATGGGAAAGATGGGAGCTCTGTCATGACAGCGCATC 1272
 Db 1093 TTCTGCCAGTGGCAGTCTGGATTAATGGGAAAGATGGGAGCTCTGTCATGACACCGCATC 1152
 Qy 1273 TTACTGACTGGTCTGGATATATGTTCTCGGAAAGATGAGCCCTGTGACACTTTGGGATTT 1332
 Db 1153 TTACTGACTGGTCTGGATATATGTTCTCGGAAAGATGAGCCCTGTGACACTTTGGGATTT 1212
 Qy 1333 GCACCCATAGTGGAAATGTGTATTAATTCAGACTGACAGATTAATGAAGATACAGT 1392
 Db 1213 GCACCCATAGTGGAAATGTGTATTAATTCAGACTGACAGATTAATGAAGATACAGT 1272
 Qy 1393 CTTGACTGCCCTTCAACATGCCCATGAGTCTGACACAACTTTGGGATTAATGAT 1452
 Db 1273 CTTGACTGCCCTTCAACATGCCCATGAGTCTGACACAACTTTGGGATTAATGAT 1332
 Qy 1453 GGAGAGGGAACATGTGTAAAGTTCGAGGGAACATCATGTCCTCACTTTGGGAG 1512
 Db 1333 GGAGAGGGAACATGTGTAAAGTTCGAGGGAACATCATGTCCTCACTTTGGGAG 1392
 Qy 1513 CCCAATGGAGTCTTCTCTCTGTCACCTCTGAGCGCGCTGATATCTACACAAATTTCTAAGC 1572

Db 1393 GCGAATGGAGTCTTCTTCTGTGCTCACCTGTGAGCGCGCCAGTATCTACACAAATTTCTAAGC 1452
 Qy 1573 ACCGCTCAAGCTATCTGCTCTGCTGATCAGCCAAAGCCTGTGAGAGGATACAGATATCTT 1632
 Db 1453 ACCGCTCAAGCTATCTGCTCTGCTGATCAGCCAAAGCCTGTGAGAGGATACAGATATCTT 1512
 Qy 1633 GAGAAATTCGAGGAAATATATGATGCAACACAGCTGAGTGTGAGTGTGCGAGAG 1692
 Db 1513 GAGAAATTCGAGGAAATATATGATGCAACACAGCTGAGTGTGAGTGTGCGAGAG 1572
 Qy 1693 AAGCCAGACTCTGCTGCTGCTTAAAGGACATCTGTAAAGCCTGTGAGTGTGCGAT 1752
 Db 1573 AAGCCAGACTCTGCTGCTGCTTAAAGGACATCTGTAAAGCCTGTGAGTGTGCGAT 1632
 Qy 1753 GGTATTTGGAAGAAATGTGAGACTTAAATTTATGTCAGCAGCAGAGGACCAATTTGTGG 1812
 Db 1633 GGTATTTGGAAGAAATGTGAGACTTAAATTTATGTCAGCAGCAGAGGACCAATTTGTGG 1692
 Qy 1813 CATGACATCTGTGTGCGGAGGAGACATGTGTGAAATATGTGTAATGATGTAAGGCCCAAGCCC 1872
 Db 1693 CATGACATCTGTGTGCGGAGGAGACATGTGTGAAATATGTGTAATGATGTAAGGCCCAAGCCC 1752
 Qy 1873 ACCCATGGCCACTGTGTGAGCTGTCTTCTGTGTCCTCATGCTCCAGGACCTGTGAGAGG 1932
 Db 1753 ACCCATGGCCACTGTGTGAGCTGTCTTCTGTGTCCTCATGCTCCAGGACCTGTGAGAGG 1812
 Qy 1933 GAGATATCTCATAGAGTGGCTCTGACCAACCCCAAGCCATCGCATGAGGAGGAAATTC 1992
 Db 1813 GAGATATCTCATAGAGTGGCTCTGACCAACCCCAAGCCATCGCATGAGGAGGAAATTC 1872
 Qy 1993 TGTGAGGGTCCACTGCACTCTGAAAGCTCTGCAACAGCTCAGAAATGTCCCGGGAGCAGT 2052
 Db 1873 TGTGAGGGTCCACTGCACTCTGAAAGCTCTGCAACAGCTCAGAAATGTCCCGGGAGCAGT 1932
 Qy 2053 GTTGACTTCGCTGTGCTCTGAGTGTGCGAGCACAACAGCAGAGATTCAGAGGGCGGCAC 2112
 Db 1933 GTTGACTTCGCTGTGCTCTGAGTGTGCGAGCACAACAGCAGAGATTCAGAGGGCGGCAC 1992
 Qy 2113 TACAGTGTGAGGCTTACACTCATAGTAGAAGTACAGACTTATGCAAACTCTACTGTATC 2172
 Db 1993 TACAGTGTGAGGCTTACACTCATAGTAGAAGTACAGACTTATGCAAACTCTACTGTATC 2052
 Qy 2173 GCGAAGGATTGATTTCTTCTTCTTGTCAAATAAAGTCAAGATGGGACTCTCATGTC 2232
 Db 2053 GCGAAGGATTGATTTCTTCTTCTTGTCAAATAAAGTCAAGATGGGACTCTCATGTC 2112
 Qy 2233 TCGAGAGATGAGCTTAATGTTGTATGAGTGTGGATGTGAGAGTGTGAGTGTGAGAT 2292
 Db 2113 TCGAGAGATGAGCTTAATGTTGTATGAGTGTGGATGTGAGAGTGTGAGTGTGAGAT 2172
 Qy 2293 GTTCTTGGATCTGATCTGTGTGAGAGCTCTGTGAGGGTGTGTACCGGATTAATCAGCC 2352
 Db 2173 GTTCTTGGATCTGATCTGTGTGAGAGCTCTGTGAGGGTGTGTACCGGATTAATCAGCC 2232
 Qy 2353 TGCAGAGTTCACAGGGCTCTCTACACAGCACAACACACACACACACACATGATATACATG 2412
 Db 2233 TGCAGAGTTCACAGGGCTCTCTACACAGCACAACACACACACACACATGATATACATG 2292
 Qy 2413 GTACCATTTCTTCTGAGCGCGGAGTATCGGATCTATGAAATGAAGTCTCTACTCTCC 2472
 Db 2293 GTACCATTTCTTCTGAGCGCGGAGTATCGGATCTATGAAATGAAGTCTCTACTCTCC 2352
 Qy 2473 TACATTTCTGTGCGCAATGCGCTCAGAGGTAATCTGATGATGSCACTGACCGTGTGAC 2532
 Db 2353 TACATTTCTGTGCGCAATGCGCTCAGAGGTAATCTGATGATGSCACTGACCGTGTGAC 2412
 Qy 2533 TGGCCCGGCGGTAACAAATTTTCTGGGCACTACTTTTCTGACTACAGAGCGTCTCTATATGAG 2592
 Db 2413 TGGCCCGGCGGTAACAAATTTTCTGGGCACTACTTTTCTGACTACAGAGCGTCTCTATATGAG 2472
 Qy 2593 CCCGAGAGATTAATTCCTACTGCGACCAACCAAGAGACATGATTTGTGGAGTCTTGTGTT 2652
 Db 2473 CCCGAGAGATTAATTCCTACTGCGACCAACCAAGAGACATGATTTGTGGAGTCTTGTGTT 2532

```

QY 2653 CAGGAGAGAAACCCGGTGTGTGCTGGGAATACCTCATGCTGCTTGGGGACGAGAG 2712
DB 2533 CAGGAGAGAAACCCGGTGTGTGCTGGGAATACCTCATGCTGCTTGGGGACGAGAG 2592
QY 2713 CAGCCCTGCTCCACGACGACGACGACGACGACGACGACGACGACGACGACGACG 2712
DB 2593 CAGCCCTGCTCCACGACGACGACGACGACGACGACGACGACGACGACGACGACG 2652
QY 2713 TGGCGAGGGG 2783
DB 2653 TGGCGAGGGG 2653

RESULT 2
US-09-930-872-5
; Sequence 5, Application US/09930872
; Patent No. 6448388
; GENERAL INFORMATION:
; APPLICANT: Fiddle, Carl Johan
; APPLICANT: Hilbun, E-zin
; TITLE OF INVENTION: No. 6448388el Human Proteases and Polynucleotides Encoding the sa
; FILE REFERENCE: LEX-0219-USA
; CURRENT APPLICATION NUMBER: US/09/930,872
; CURRENT FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: US 60/225,852
; PRIOR FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq For Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 4042
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-930-872-5

Query Match 54.8%; Score 1874.6; DB 4; Length 4042;
Best Local Similarity 84.4%; Pred. No. 0;
Matches 2287; Conservative 0; Mismatches 179; Indels 245; Gaps 6;

QY 202 CGAGGACCTCGTGGCGACGACGACGACGACGACGACGACGACGACGACGACGACG 261
DB 167 CGAGGACCTCGTGGCGACGACGACGACGACGACGACGACGACGACGACGACGACG 226
QY 262 CGTGTCTCTCCACCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 321
DB 227 CGTGTCTCTCCACCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 273
QY 322 GCTGGAAACCATCGTCTCAGCAAACTAACACAGGAAACAGAAACCAACACACTGCATGTT 381
DB 274 ----- 273
QY 382 CTCACATATGACCTGCTCTGCTACAGGTTGACACAGGAGCGGATACGTTGTC 441
DB 274 -----AAATGACCTGCTCTGCTCTACAGGTTGACACAGGAGCGGATACGTTGTC 326
QY 442 CATGAATCATGCACCATCAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 501
DB 327 CATGAATCATGCACCATCAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 386
QY 502 CTCACCTTCGGCTGAAGAGGCCGCCAGCA CAGCTTCCATGATCTGAGGACTTCGAGC 561
DB 387 CTTCACTTCGGCTGAAGAGGCCGCCAGCA CAGCTTCCATGATCTGAGGACTTCGAGC 446
QY 562 AGCTAGTGGCTCTGCTTTATGTCGAGAGCTTGGGAAGGACAGGACATAGTCTGTG 621
DB 447 AGCTAGTGGCTCTGCTTTATGTCGAGAGCTTGGGAAGGACAGGACATAGTCTGTG 506
QY 622 CAGCATCTCCGCGAGGACCTCTCTTCTCTCAGAGGCTCTTGGATCAGGAGGAGGAC 681
DB 507 CAGCATCTCCGCGAGGACCTCTCTTCTCTCAGAGGCTCTTGGATCAGGAGGAGGAC 566
QY 582 TGGCCATCGCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 741
DB 582 TGGCCATCGCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG

```

QY	1693	AAAGCCAGACTCTGCGATGCTGGACCTTTAAAGAGGACATCTGTAAAGCCCTGTGTGGCCAT	1752
QY	1694	AAAGCCAGACTCTGCGATGCTGGACCTTTAAAGAGGACATCTGTAAAGCCCTGTGTGGCCAT	1753
DB	1671	AAAGCCAGACTCTGCGATGCTGGACCTTTAAAGAGGACATCTGTAAAGCCCTGTGTGGCCAT	1730
QY	1753	CGATTGGAGAGAAATGTGAGACCTAAATTTATGCGACGACGACGAAGGCGCAATTTGTGGG	1812
QY	1731	CGATTGGAGAGAAATGTGAGACCTAAATTTATGCGACGACGACGAAGGCGCAATTTGTGGG	1790
QY	1813	CATGACATCTGGTCCGGGAGGACAGTGTGCGAATATGCGATGATGACGAGCCCGCAGGCC	1872
DB	1791	CATGACATCTGGTCCGGGAGGACAGTGTGCGAATATGCGATGATGACGAGCCCGCAGGCC	1850
QY	1873	ACCACATGCGCACTGGTGGGACCTGGTCTCTTGTGGTCCCATGCTCCAGAGACTCGCGAGGG	1932
DB	1851	ACCACATGCGCACTGGTGGGACCTGGTCTCTTGTGGTCCCATGCTCCAGAGACTCGCGAGGG	1910
QY	1933	GGAGTATCTCATAGAGAGTCGGCTCTGCAACCAACCCGACGACCATCGCATGAGAGGAAGTTC	1992
DB	1911	GGAGTATCTCATAGAGAGTCGGCTCTGCAACCAACCCGACGACCATCGCATGAGAGGAAGTTC	1970
QY	1993	TGTGAGGGCTCCACTCGCACTCTGAAGCTCTGCAACGATGAGAAATGTCCCGCGACAGT	2052
DB	1971	TGTGAGGGCTCCACTCGCACTCTGAAGCTCTGCAACGATGAGAAATGTCCCGCGACAGT	2030
QY	2053	GTTGACTTCCGGTGTCTCAGTGTGGCCGAGACCAACGACGACGATTTCAGAGGGCGGAC	2112
DB	2031	GTTGACTTCCGGTGTCTCAGTGTGGCCGAGACCAACGACGACGATTTCAGAGGGCGGAC	2090
QY	2113	TACAGGTGGAAGCCTTACCTCAAGTAGAAGACTCAGGACTTATGCAAACTCTACTGTATC	2172
DB	2091	TACAGGTGGAAGCCTTACACTCAAGTAGAAGACTCAGGACTTATGCAAACTCTACTGTATC	2150
QY	2173	GCAGAGGATTTGATTTCTTTCTTTTGTGCAAAATAAAGTCAAAAGTGGGACTCCATGC	2232
DB	2151	GCAGAGGATTTGATTTCTTTCTTTTGTGCAAAATAAAGTCAAAAGTGGGACTCCATGC	2210
QY	2233	TGGAGGATAGCCGTAATTTGTATAGATGGGATATGTGAGAGAGTGGATGTGTAACAT	2292
DB	2211	TGGAGGATAGCCGTAATTTGTATAGATGGGATATGTGAGAGAGTGGATGTGTAACAT	2270
QY	2293	GTCTTGGATCTGATGCTGTGGAAGAGCTGTGGGGTGTTTAAACGGGAATACTCAGCC	2352
DB	2271	GTCTTGGATCTGATGCTGTGGAAGAGCTGTGGGGTGTTTAAACGGGAATACTCAGCC	2330
QY	2353	TGCGAGATTCACGGGGTCTTACACGAGACCAACGACACGACGATATATCATG	2412
DB	2331	TGCGAGATTCACGGGGTCTTACACGAGACCAACGACACGATATATCATG	2390
QY	2413	GTGACATCTCTCTGTGAGCCGGAGTATCGGATCTATGAAATGGAAGCTCTTACCTCC	2472
DB	2391	GTGACATCTCTCTGTGAGCCGGAGTATCGGATCTATGAAATGGAAGCTCTTACCTCC	2450
QY	2473	TACATTTCTGTGGCAATGCTCCAGAGGACTACTCTGAATGGGCACTGGACCTGGAC	2532
DB	2451	TACATTTCTGTGGCAATGCTCCAGAGGACTACTCTGAATGGGCACTGGACCTGGAC	2510
QY	2533	TGGCCCGGCGGTACAAAATTTTTCGGGCACTACTTTCGATACAGAGGTCCTATATATGAG	2592
DB	2511	TGGCCCGGCGGTACAAAATTTTTCGGGCACTACTTTCGATACAGAGGTCCTATATATGAG	2570
QY	2593	CCGAGAACTTAATGCTACTGAGACCAACCAAGAGACACTGATTTGGAGGCTGCTGTTT	2652
DB	2571	CCGAGAACTTAATGCTACTGAGACCAACCAAGAGACACTGATTTGGAGGCTGCTGTTT	2630
QY	2653	CAGGGAAGGAAACCCGGGTGTGCTGGGAATACTCCATGCTCTTGGGACCGAGAG	2712
DB	2631	CAGGGAAGGAAACCCGGGTGTGCTGGGAATACTCCATGCTCTTGGGACCGAGAG	2690
QY	2713	CAGCCCCCTGCCAGGCCAGCTTACACTTGGGCGCATGTGGGCTCTGAGTGCTCGGTGCC	2772
DB	2691	CAGCCCCCTGCCAGGCCAGCTTACACTTGGGCGCATGTGGGCTCTGAGTGCTCGGTGCC	2750

```

QY 2773 TCGGAGGGGG 2783
DB 2751 TCGGAGGGGG 2761

RESULT 3
US-09-930-872-1
/ Sequences 1, Application US/09930872
/ Path No. 6448349
/ GENERAL INFORMATION:
/ APPLICANT: Priddle, Carl Johan
/ APPLICANT: Hilpun, Erin
/ TITLE OF INVENTION: No. 6448388el Human Proteases and Polynucleotides Encoding the S
/ PTLE REFERENCE: LEX-0219-USA
/ CURRENT APPLICATION NUMBER: US/09/930.872
/ PRIOR FILING DATE: 2001-08-14
/ PRIOR FILING DATE: 2000-08-16
/ NUMBER OF SEQ ID NOS: 5
/ SOFTWARE: FASTSEQ for Windows Version 4.0
/ SEQ ID NO 1
/ LENGTH: 1476
/ TYPE: DNA
/ ORGANISM: homo sapiens
US-09-930-872-1

Query Match 22.9%; Score 663; DB 4; Length 1476;
Best Local Similarity 71.5%; Pred. No.1.4e-180;
Matches 1079; Conservative 0; Mismatches 185; Indels 245; Gaps 6;

QY 202 GCAGGCACTGCTGCTGCGCCATGGGACCCGCGACGGCAGCGCTGGAGAGCCCGACGCTGCC 261
DB 69 GCAGGCACTGCTGCTGCGCCATGGGACCCGCGACGGCAGCGCTGGAGAGCCCGACGCTGCC 128
QY 262 GCGTCTCTCTCCACCCGCGAGCGCGCTGGATGGATGGAAGGCGCGACATGGATGAA 321
DB 129 GCGTCTCTCTCCACCCGCGAGCGCGCTGGATGGATGGAAGGCGCGCGG 175
QY 322 GCTGGAACCACTCGTTCTCAGCACTAAACACAGGAACAGAAAAACCAACACTGCATGTT 381
DB 176 ----- 175
QY 382 CTCACCTCAATATGACCTGGTCTCTCCCTACGAGGTTGACACAGGGCGATTAGGTGTC 441
DB 176 -----ANTATGACCTGGTCTCTCCCTACGAGGTTGACACAGGGCGCGATTAGGTGTC 228
QY 442 CATGAATATATGCAATCATGACGCGGCGGAGAGACGATGGCGTTCGAGAGTGTAGTCT 501
DB 229 CATGAATATATGCAATCATGACGCGGCGGAGAGACGATGGCGTTCGAGAGTGTAGTCT 288
QY 502 CTTGACCTTGCGCTGAAAGCGACGACGATCTTCACTGATCTGAGACTTCGACG 561
DB 289 CTTGACCTTGCGCTGAAAGCGCTTCAAGCGACCTTCACGTGATCTGAGACTTCGACG 348
QY 562 AGCCTAGTGGCTCTTGGCTTTATTCGACAGCTTGGGAAAGACAGGCACTAGTCTGTG 621
DB 349 AGCCTAGTGGCTCTTGGCTTTATTCGACAGCTTGGGAAAGACAGGCACTAGTCTGTG 408
QY 622 CAGACTTTTACCGCAGAGGACTCTGTTCTTATCATAGCTCTTTGCCATCAAGAAAC 681
DB 409 CAGACTTTTACCGCAGAGGACTCTGTTCTTATCATAGGCTCTTTGCCATCAAGAAAC 468
QY 682 TCGCATATGCATGGAGGGAGTCTGTGTGAGGCGCTCCACTGCGACTCTTGAGCTTGCAC 741
DB 469 TCTTCAGTGG -----CCCTTCACCTGCCAGGCTTGCA 504
QY 742 AGTCAGAAATGTCCTCCCGGACAGCTGTGATCTGCTGCTGCTCAGTGTGCGGACGAC 801
DB 505 GGCATGATACGACAGAGAGGACGATTACTTCTTACGCTGCTTCTTCACTCTCA 564
QY 802 AGCAGACGATTCAGAGGCGCGCATACAGTGG --AAGCTTTACACTAGTAGAGCG 859
DB 565 TGAAGCTTCGCGAGAGCTGCGCCAAAGGAGGCTCTGCCATCTTCACTAGTACAGAGATCC 624

```

QY	860	ACTATATGCGAACTCTACTATATGCCAGAGCAATTGATTTCTTCTTTCTTTGTCAAATA	919
DB	635	ACAGGCCCATGCTCTGGGCGAGTAGAGCTCTGGTCACTC-----AAGG	672
QY	920	AAGTCAAGATCGAGATCTGCTGCTCGAGGATAGCGTAATGTTTGTATAGATGGGATAT	979
DB	673	ACATGGGAGCTGGCAATCAACCCCTGCACAGACCTTGCCTGGGACTGCCACAA	732
QY	980	TGAGCTCAGTGGTGTCCACATCTGGCC--ACAATGCCAGCTCCCAAGAGAGACCTC	1038
DB	793	AAAGCGCATTTCTGTGGAGAGCGCAGAAATACATGCCAGGCTCTCTTGAGAGACTC	792
QY	1039	TTTCACTTCCGAGATGAGTATAGTCTTGTCTTAGCGGATAGCGGCTCTCTCTGAGTCC	1098
DB	793	TTTCACTTCCGAGATGAGTATAGTCTTGTCTTAGCGGATAGCGGCTCTCTCTGAGTCC	852
QY	1099	CATAGAAATGAGAACTGAACGTGGAGACCTTGGTGGTCTGCACAAAAGACATGATGCAA	1158
DB	853	CATAGAAATGAGAACTGAACGTGGAGACCTTGGTGGTCTGCACAAAAGACATGATGCAA	912
QY	1159	AACCTGGCCGATGAAATATCAACCATACGTGCTCAAGATCTCAACATGGTATCTGCT	1218
DB	913	AACCTGGCCGATGAAATATCAACCATACGTGCTCAAGATCTCAACATGGTATCTGCT	972
QY	1219	TTATTCCAAAGA-----	1229
DB	973	TTATTCCAAAGATGGAACTAGGAGGAACATCAACATTGCATTTAGGCTGATCTTT	1032
QY	1230	-----	1239
DB	1033	CTAGAGATGACAGCCAGCATGTTGATAAGTCAACCGACGAGACCAACCTTAAGTAGC	1092
QY	1230	-----TGGAATCATGGGAGAAAGTGGGACTCTCATGACGACGCGCATC	1272
DB	1093	TTCTGCCACTGGCAGTCTGGATTTGATGGGAGAAAGTGGGACTCTCATGACGACGCGCATC	1152
QY	1273	TTACTGACTGGTCTGATATATGTTCTTGGAGAAATGACCTCTGACACATTTGGGATTT	1332
DB	1153	TTACTGACTGGTCTGATATATGTTCTTGGAGAAATGAGGCTCTGACACATTTGGGATTT	1212
QY	1333	GCACCATGAATGGAAATGATGATGAATATGCGAGTGCACAGTAAATGAGAGATACNGT	1392
DB	1213	GCACCCATGAATGGAAATGATGATGAATATGCGAGTGCACAGTAAATGAGAGATACNGT	1272
QY	1393	CTTGGACCTGGCTTCAACATTCGCATGAGTCTGCACACACTTTGGCATGATTTCAATGAT	1452
DB	1273	CTTGGACTGGCTTCAACATTCGCATGAGTCTGCACACACTTTGGCATGATTTCAATGAT	1332
QY	1453	GGAGAGGGAAATGTTAAAAGTCCGAGGAGCAACATCATGCTCCCTACATCTGGCAGGA	1512
DB	1333	GGAGAGGGAAATGTTAAAAGTCCGAGGAGCAACATCATGCTCCCTACATCTGGCAGGA	1392
QY	1513	GGCAATGGAGTCTTCTTGGTTCACCTCGACGCGCGAGTATCTACACAAATTTCTAAGC	1572
DB	1393	GGCAATGGAGTCTTCTTGGTTCACCTCGACGCGCGAGTATCTACACAAATTTCTAAGC	1452
QY	1573	ACCGCTCAA	1581
DB	1453	TCAGTGAAT	1461

RESISTANCE

RESIDUE 4	1912	TGCTCCAGGACCTCTGGAGGGGGAGTATCTATNAGGAGTGCCTCTTGCAACACCCCAAG	1971
US-09-392,184-5/c			
Publication US/09392184			
Sequence 5, 355589			
Sequence 6, 355589			
GENERAL INFORMATION:			
APPLICANT: Robison, Keith E.			
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN			
TITLE OF INVENTION: PROTEASE HOMOLOGS			
FILE REFERENCE: 5800-55			
CURRENT APPLICATION NUMBER: US/09/392,184			
CURRENT FILING DATE: 1999-09-09			

3550	Qy	DB	CAGGCTCGCTGCTCGCGCCCTCTCTTCGCCACGTCGAGTCGACCACTTTGAGCTT	3491
2092	Qy	DB	AGAGGATTCAGGCGCGCATCAGATGAGGAGCCCTTACATCACTCAGTAGAGATCAGAC	2151
3490	Qy	DB	ATGCTCTACAGAGCGCAGCTGCACACATGGGTGC-----CGTGGTCAATGACGTGAC	3437
2152	Qy	DB	TTATGCAAACTTACTGTATCGACAGAGGATTTGATTTCTCTTTCTTTGTCAATAAA	2211
3436	Qy	DB	CCCTGCGAGCTGCACCTGCGCGCCGCGAATGAGTACTTTGCCGAGAAGCTCGGGACGC	3377
2212	Qy	DB	GTCAAAAGATGGGACTCCATGCTCGGAGGAT-----AGCCGTAAATGTTGTATAGATGG	2265
3376	Qy	DB	GTGTCGATGTCACCCCTGCTTACAGGTCCGAGCCAGCCGGGACCTCTGCATCAACGC	3317
2266	Qy	DB	MTATGTGAGAGAGTTGGATGTGCAATGTCTTGGATCTGATGCTGTGAAGACGCTCTGT	2325
3316	Qy	DB	ATCTGTAAAGACGTGGGCTGTGACTTCSAGATTTGATCTCCGGTGTATGAGAGCAACGCTGT	3257
2326	Qy	DB	GGGCTGTGTAACCGGAAATAACTCAGCTTCGACAGTTTACAGGGGCTCTTACACCAAGAC	2385
3256	Qy	DB	GGTGTGTGCCACGCAACCGCTCCACCTGCACACCGTGACGGGACCTTTGC---AGGAG	3200
2386	Qy	DB	CACCACCAACCAAGTATTATACATGTCACCAATTCTCTTGAGCGCCGAGATTCGCG	2445
3199	Qy	DB	CCGAGGGCTGGGGTATGTGATGTGTGGGCTGATCCAGCGCGCGACGACGAGATTCGC	3140
2446	Qy	DB	ATCTATGAAATGAACGCTCTACCTCTCTACATTTCTTGTGCCA---ATGCCCTCAGAAG	2502
3139	Qy	DB	ATCCAGAGAGGTGCCGAGGCTGCCAATCTCTGCGACTCTCGGAGTGAGAGCCCGGAGAG	3080
2503	Qy	DB	TACTACTGAATGGCACTTGACCGTGGAGTGGCGCGCGGCTACAAATTTTCGGGCACT	2562
3079	Qy	DB	TACTTCTCAATGTGGCTTGACGACATCCAGTGGAGAGCGGGACTACCAAGGTGCGAGGGCC	3020
2563	Qy	DB	ACTTTGCGCTACAGACGGTCTTATATGAGCCCGAGACTTAATCGGTACTGAGCAACCC	2622
3019	Qy	DB	ACCTTCATCATACGACGAGGGGCACTGG---GAGACCTCGTCCCCGGTCCGCTCC	2963
2623	Qy	DB	AACGAGACACTGATTTGGAGCTCTGTGTTTACGAGAGAGAACCGGGTGTGCTCTGGGAA	2682
2962	Qy	DB	ANGCAGCTCTCTGGATTCAGCTCTGTTCACGAGAGCAACCTCGGGTGCATCAGG	2903
2683	Qy	DB	TACTCAACGCTCTGGGACCGAGAGAGAGCCCGCTCGGACCGAGTGCATCACTGG	2742
2902	Qy	DB	TTACACATCCAGAGGACGAGGTGGCCACGACGAGGTCCGCGCCGCTTCTCTGG	2843
2743	Qy	DB	GCAATCGTGGCT---CTGAGTGTCTCGGTCTCGGAGGGGCT	2784
2842	Qy	DB	CATTATGGGCTCTGGACCAAGTGTGCACAGTCACTTGGCGCAGAGGT	2798

DB	Accession	Length	Sequence
Db	US-02-369-364A-6	1660	TGCTCAGCGAGCTGTGCATGGCGTACAGAGCGCGAGCGGCACTGCACGCGCTTACG 1711
Qy	Sequence 6, Application US/09369364A	1972	CCATCGCATGGAGGGAAGTTCTGTGGGSGCTCCACTCGCACTCTGAAGCTCTGCGAAGCT 2033
Db	GENERAL INFORMATION:	1720	CCCAAATACAAAGGCAGATACTGTGTGGTGGAGCGCAAGCGCTTCGCGCTCTCTGCACCTG 1779
Qy	APPLICANT: Aptec, Suneel	2032	CAGAAATGTCCTCCCGGGAAGTGTTCACCTTCGTCGTCTCAGTGTGGCGGAGCAACAAGC 2093
Db	APPLICANT: Hurskainen, Tiina L.	1780	CAGGCTGTGGCGCTGTGGCGCGCCCTCTCTCCGCACATCTCGTCGACGACTTTGACGCT 1833
Qy	APPLICANT: Hirohata, Satoshi	1980	AGACGCTCTGAGGCGGCGACCTACAGCTGGAAAGCTTACATCAAGTGAAGTACAGCAGC 2155
Db	TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases	1840	ATGCTCTACAGAGCCGACGTGCACACATCTGGTGC ----- CCGTGCTCATGATCGTGAAC 1899
Qy	CURRENT APPLICATION NUMBER: US/09/369,364A	2152	TTATGCAAACTCTACTGTATCGCAGAGGATTTGATTTCTTTCTTTTGTTCGCAATAAA 2211
Db	CURRENT FILING DATE: 1999-08-06	1894	CCCTGCGAGCTGCATCGCGCGCCCGCAATGATGACTTTTGCACAAGAGCTGCGGGAAGCC 1953
Qy	NUMBER OF SEQ ID NOS: 31	2212	GTCAAAGATGGGACTTCCATGCTGGAG ----- GATAGCGGTAATGTTGTATGATGGG 2265
Db	SOFTWARE: Patentin Ver. 2.1	1954	TGTGTGCTGTGGCAACCCCTTGCTACCGGTGTCGAGCGCGGAGCTCTTGGATCAACGGC 2013
Qy	SEQ ID NO 6		
Db	LENGTH: 3218		
Qy	TYPE: DNA		
Db	ORGANISM: Homo sapiens		
Qy	FEATURES: ADAMTS-7		
Db	FEATURES: CDS		
Qy	NAME: X		
Db	LOCATION: (13)...		
Qy	US-09-369-364A-6		

```
QY 2266 ATATGTGAGAGGTTGGATGTGCAATGCTCTTGAGTCTGATGCTGTCTGAGAGGTCGT 2325
Db 2014 ATCTGTAAAGACATGGGCTGTGTGACTTCAGATTTGACTCCGGTGTATGGAGCGCGTGT 2073
QY 2326 GGGGCTGTGTAACGGGGAATACCTCAGCTGCGACAGATTCACAGGGGTCTCTACACCAAGAC 2385
Db 2074 GGTGTGTGCCACAGGCAAGGCTCCACCTGGCCACACACCGTGTGAGCGGGACCT---TCGAGGAG 2130
QY 2386 CACCACACCAACCAAGTATTATACATGAGTGTACCAATTCCTCTGSGAGCCGGAGATTCGCG 2445
Db 2131 GCCAGGGTCTGGGGTATGTGTGATGTGTGGGCTGATCCAGAGCGGGCGACGGAGATCCGC 2190
QY 2446 ATCTATGAATGAACGTCTTACCTCTACATTTCTGTGC---GCAATGCCCTCAGAGG 2502
Db 2191 ATCCAGAGGTTCGCCAGGCTGCCAACTTCCTGCGCACTGCGGAGCGGAGCCCGGAGAG 2250
QY 2503 TACTACTGTAATGGGCACTGAGCGGTGAGCTGGCGCCGGCCGCTACAAATTTTCGGGCACT 2562
Db 2251 TACTTCTCAATGTGGCTGGACCATCAGTGTGAACCGGGACCTACCAAGTGTGCGAGGAC 2310
QY 2563 ACTTTTCGACTACAGAGGCTCTATATAGAGCCCGAGAACTTAATCTGCTACTGGAACCAAC 2622
Db 2311 ACCTTTCACATACGACGAGCGGCGCACTGG---GAGAACTCTACGTCCTCCCGGGTCCAC 2367
QY 2623 AACGAGACACTGATTTGTGAGCTGC 2647
Db 2368 AAGGAGCCTGTCTGGATCCAGGTGC 2392

RESULT 6
US-09-369-364A-16
; Sequence 16, Application US/09369364A
; Patent No. 6391610
; GENERAL INFORMATION:
; APPLICANT: Apte, Suneel
; APPLICANT: Hurskainen, Tiina L.
; APPLICANT: Hirohata, Satoshi
; FILE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
; TITLE REFERENCES: 26473/4007/10-30-00
; CURRENT APPLICATION NUMBER: US/09/369,364A
; CURRENT FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 3885
; TYPE: DNA
; ORGANISM: Homo sapiens ADAMTS-10
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3246)
; NAME/KEY: misc.feature
; LOCATION: (3877)
; OTHER INFORMATION: n=A
; NAME/KEY: misc.feature
; LOCATION: (3882)
; OTHER INFORMATION: n=C
US-09-369-364A-16

Query Match 6.1%; Score 176.8; DB 4; Length 3885;
Best Local Similarity 49.4%; Pred. No. 2e-40;
Matches 705; Conservative 0; Mismatches 687; Indels 36; Gaps 8;

QY 1240 GGGAAAGATGGGACTCGTCATGACACACGCCATCTTACTGACTGGTCTGGATATATGTTC 1299
Db 931 GAGAACGGTGTGGCTAAACATGACACAGCAGTGTCTCATCACAGCTGTATGATCTGCATC 990
QY 1300 TGGAAAGATGAGCCCTGTGACACTTTTGGGATTTGCAACCCATTAAGTGGAAATGTGTAGTAA 1359
Db 991 TACAAAGACAAACCCCTGGGACACTAGGCTGGC---CCGGTGGGCGGAATGTGTGAGGCG 1049
QY 1360 TATCGACCTGCACAGATTAATGAAGATACAGTCTTTGAGCTGGCCCTTCACCATTTGCCAT 1419
Db 1050 GAGAGAGCTGCACGCTCAATGAGGACATTGG---CTGCCACAGCGTTTCACCATTG---CCAC 1107

1420 GAGTCTGACACAACTTTGGCATGATTCATGATGGAGAGGGAACATGT-----GT 1470
1108 GAGTCTGGGACACAACTTTGGCATGATTCATGATGGAGAGGGAACATGTGGGGGCGCGT 1167
1471 AAAAGTCTGGAGGACACATGATGCTCCCTACATTTGGAGAGCGCAATGGAGTCTTCTCC 1530
1168 GGTGAGGACCGACGACGATGCTGCTGCCCATTTACATGAGAGACCAACCATTTGCG 1227
1531 TGGTCACTCTGAGCGGCGGAGTATCTACAAATTTCTAAGCAACGCTCAAGCTATCTGCG 1590
1228 TGGTCACTCTGAGCGGCGGAGTATCTACAAATTTCTAAGCAACGCTCAAGCTATCTGCG 1287
1591 CTTGCTGATCAGCGGCGGAGTATCTACAAATTTCTAAGCAACGCTCAAGCTATCTGCG 1650
1288 CTGACAAACCGGCGGAGTATCTACAAATTTCTAAGCAACGCTCAAGCTATCTGCG 1344
1651 TTATGATCTGACAAACGAGTATCTACAAATTTCTAAGCAACGCTCAAGCTATCTGCG 1710
1345 GCTTACGATGAGTATGAGCAATGCTCTTACATGAGGATCAATTCGCTGAGTGTAA 1404
1711 CTGGAATTTAAAGGAGTATCTGTAAGCTCTGCTGCTCATCTGTTTGGAGAGAAATGT 1770
1405 TAGCG-----GGAGTCTGCGAGGCTGTGTGTCTGAGAGAGGACACCGGTCG 1455
1771 GAGATCAATTTTATGCGAGCGACGAGGAGCAATTTGTGG-----CATGACATG 1821
1456 ATCACCAACAGCATCCCGCGCGGCGACCTGTGTCNAGAGCACACCATCTGACAG 1515
1822 TGGTCTGGGCGGAGGAGTGTGTGAATTTGTGATGAGAGGCGCCCAAGCCCAACCATGCG 1881
1516 GGTGTGTGCTACAAAGGCTGTGTGCTCCCTTTGGGTGCGCGCCAGAGGGTGTGACCG 1575
1882 CACTGCTGAGAGTGTCTTCTGCTCCCATGCTTCAGAGACTGCGGAGGCGGAGTATCT 1941
1576 GCTTGGGCGGCGTGTGATCTCTGCGGCGCATGTCGCGCGGAGCTGTGGCGGCGGTCG 1635
1942 CATAGGATGCGCTCTGACAAACCCCAAGCGATGCMATGAGGAGGAGTCTCTGAGGCG 2001
1636 TCTTCTAGTGTCTGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1695
2002 TCCATCTGCACTGTGAGCTCTGCAACAGTCAAGATGTCCTCCGCGGAGAGTGTGACTTC 2061
1696 GAGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1755
2062 GGTGCTGCTGAGTGTGCGAGACACACAGAGAGCGATTCAGAGGCGGCGGCGGCGGCGG 2121
1756 AGAGAGTGTGCTGTGATTTGAGAGTCTGAGAGTCTGAGAGTCTGAGAGTCTGAGAG 1815
2122 AGCGCTTACACTCAGTATGAGATCAGGATTTATGCAAACTTACTGTATGCGAGAGGA 2181
1816 AAAAGTACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1872
2182 TTTGATTTCTCTTTTCTTTTCTGCAATTAAGTCAAGATGCGAGCTCCATGCTCGAGGAT 2241
1873 TTTCACTTTTACACGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1932
2242 AGCGGTAATTTGATGATGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 2301
1933 AGGTTGAGCATTTTGGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1992
2302 TCTGATCTGTTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2361
1993 TCCGACTCTGGGAGGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2052
2362 CACAGGGTCTCTACACAGACCAACACACACAGATTTATTCATGAGTGTGAGTGTGAG 2421
2053 ATGAGGGGCTCTTACACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2112
2422 CTTCTGAGAGCGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2481
2113 CCCAAGGCTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2172
```

Qy 2482 GTGGCGCATGCCCTCAGAGAGTATACCTGAATGGGACATGGACCGGTGACGTGGCCCGCC 2541
 Db 2173 CTGAAGGGAGCAGAGAGTCCCTGCTGCTGGAGGGCTGCTGGGACCCCGCCGCCAC 2322
 Qy 2542 CGGTACAAATTTTGGGCACTACTTTCGACTACAGAGGTCTCTATATGCGCGGAGAC 2601
 Db 2233 CGTCTGCTCTAGCTGGGACCACTTTCACCTGCGACAGGGCCACACGCTCGAGCC 2292
 Qy 2602 TTATGCTGCTACTGACCAACCAACGACGACACTGATTTGGAGCTGCTG 2649
 Db 2293 CTCGAAGCCCTGGACCGGATTAATGTCATCTCTCATGCTGCTGGTGGCTG 2340

RESULT 7
 US-09-369-364A-4
 ; Sequence 4, Application US/09369364A
 ; Patent No. 6391610
 ; GENERAL INFORMATION:
 ; APPLICANT: Aptek, Suneel
 ; APPLICANT: Hurskainen, Tiina L.
 ; APPLICANT: Hirschbach, Satoshi
 ; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
 ; FILE REFERENCE: 26473/4007/10-30-00
 ; CURRENT APPLICATION NUMBER: US/09/369,364A
 ; CURRENT FILING DATE: 1999-08-06
 ; NUMBER OF SEQ ID NOS: 31
 ; SOFTWARE: Patent in Ver. 2.1
 ; SEQ ID NO 4
 ; LENGTH: 2848
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens ADAMTS-6
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (22)..(2601)
 ; NAME/KEY: misc feature
 ; LOCATION: (1369)
 ; OTHER INFORMATION: n = c
 ; NAME/KEY: misc feature
 ; LOCATION: (1620)
 ; OTHER INFORMATION: n=c
 US-09-369-364A-4

Query Match 5.0%; Score 144; DB 4; Length 2848;
 Best Local Similarity 48.5%; Pred. No. 4.5e-31;
 Matches 464; Conservative 0; Mismatches 481; Indels 12; Gaps 2;

Qy 1726 GACATCTGTAAGCCCTGTGGCCATCTGATTTGGAGGAAATGTGAGCTAAATTTATG 1785
 Db 1390 GAGTGTGTAGAGCTCTGTGTCTCAGCAAGACACCGTGTGTCCACACAGATTT 1449
 Qy 1786 CCAGCAGCAGGACCAATTTGGGCTGATGTGGCTGGCGGAGGACAGTGTG 1845
 Db 1450 CCAGCAGCTGAGGGGACACTGTGTCTCAACATGGGAATATTGAAGAAGGTGTTCAG 1509
 Qy 1846 AAATATGTCAT-----GAGGCGCCCAAGCCACCATGACGACCTGTGGGCTG 1896
 Db 1510 GGAGATGTGTTCTCTTTGGCATTTGSCCCAGACATAGATGGGGCTGGGCTCCCTGG 1569
 Qy 1897 TCTTCTGTGTCCTCATCTCCAGGACCTGGGAGGGGAGTATCTCATAGAGTGGCCTC 1956
 Db 1570 TCACTATGGGGAGTGTGACGAGGACCTGGGGGGAGGCTGTTCTCATCTCCTAAGACAC 1629
 Qy 1957 TGCACCAACCCNAGCCATCGCATGGAGGGAATTTCTGTGAGGGCTCCACTGCGCACTGTG 2016
 Db 1630 TGTGACAGTCCAGCACCTTCGGAGTGTGAATAATATTGCTTTGGGAAAGAAACGGTAT 1689
 Qy 2017 AAGCTCTGCAACAGTCAGAAATGTCCCGGACAGTGTGACTTCCGTGCTGCTCAGTGT 2076
 Db 1690 CGTCTGCTGAACACAGATTCATGCTCCCTTTGGGTTCGGGAGATTTTCAGAGAAACAGTGT 1749
 Qy 2077 GCGAGACCAACAGACAGACATTCAGAGGGCGGCACATCAAGTGAAGCTTTACACTCAA 2136
 Db 1750 GCAGACTTTCACATATGCTTTCCGAGGAAAGTATTATTAACATGGAACCCCTACTACT-- 1806

Qy 2137 GTAGAGATCAGGACTTATGCAAACTCTACTGTATCCGACAGGAGTTGATTTCTCTTT 2196
 Db 1907 GAGGTGGGTAACACCTTGTGCATTAACCTGCTTGGCTGAAGTATTAATTTCTACCT 1866
 Qy 2197 TCTTTGTCAATTAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 2256
 Db 1867 GACGTCTCTCTGGTGTGATGAGGACCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 1926
 Qy 2257 ATAGATGGGATGTGAGAGAGTGGATGTGACATGTCTTTGGATCTGATCTGTTGAA 2316
 Db 1927 ATCAATGGAGATGCMAGCACCTAGGCTGTGATTAATTTTGGATCTGATCTAGGAA 1986
 Qy 2317 GAGCTGTGGGGTGTGTAACGGGAATACTCAGCTCGACATTCACAGGGGTCTCTAC 2376
 Db 1987 GATAGATGTCGATCTGTGGAGGGGGGAGACATGTGATGCCATTTGAAGGGTCTTTC 2046
 Qy 2377 ACBAAGCACCCACACCAACACAGTATTATCATCATGTGTGATGACATACCAAGAGCTCTGT 2106
 Db 2047 AATGATTCACCTCCCGGAGGCTTACATGGAAGTGTGATGACATACCAAGAGCTCTGT 2106
 Qy 2437 AGTATCGGCATCTATGAATGAACCTCTCTACCTCTACATTTCTGTGGCAATGCCCTC 2496
 Db 2107 CACATGGAAGTTAGAGAGTGTCCATGTCAAGAACTATATTGCTTTAAATCTGAAGGA 2166
 Qy 2497 AGAAGTACTACTGAATGGGCACTGGACCGTGGAGCTGGCCGGCCGCTACAAATTTTC 2556
 Db 2167 GATGATTAATTAATGCTGCTGAGCTATTGAGCTGAGCTAGGAAATTTGATGTTGCT 2226
 Qy 2557 GGCATCTTTTGGACTACAGACGCTCTTAATGAGCCGAGAACTTTAATCTACTCTGGA 2616
 Db 2227 GGGACAGCTTTTCATTACAGAGACCACTGATGACCAAGATCTTTGGAAGCTCTAGGT 2286
 Qy 2617 CCACCAACAGACACTGATTTGGAGCTGCTGTTTCAGGGAGAACCCGGGTGTT 2673
 Db 2287 CCTACTCAGAAATCTCATGCTCATGCTGCTTCTTCAAGAACAGAAATTTGGGAAAT 2343

RESULT 8
 US-09-491-522-2
 ; Sequence 2, Application US/09491522
 ; Patent No. 6428998
 ; GENERAL INFORMATION:
 ; APPLICANT: College, Alain
 ; APPLICANT: Leprieux, Charles M.
 ; APPLICANT: Frockop, Darwin J.
 ; TITLE OF INVENTION: RECOMBINANT N-PROTEINASE
 ; NUMBER OF INVENTIONS: 1
 ; NUMBER OF SEQUENCES: 1
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds, LLP
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: NY
 ; COUNTRY: USA
 ; ZIP: 10036-2811
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: Windows
 ; SOFTWARE: FASTSEQ for Windows Version 2.0b
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/491,522
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/886,333
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Abrams, Samuel B
 ; REGISTRATION NUMBER: 30,605
 ; REFERENCE/DOCKET NUMBER: 8369-0060-999
 ; TELECOMMUNICATION INFORMATION:


```

, TELEPHONE: 650-493-4935
,
, TELEFAX: 650-493-8556
,
, TELEX: 66141 PENNIE
,
, INFORMATION FOR SEQ ID NO: 2:
,
, SEQUENCE CHARACTERISTICS:
,
,   LENGTH: 2450 base pairs
,   TYPE: nucleic acid
,   STRANDEDNESS: single
,   TOPOLOGY: linear
,
US-09-491-522-2

```


QY 1308 TGAAGCTGCTGAGACATTTGGGATTTGACCCATTAAGTGGATGTGTAAATATGCGAG 1367
 Db 1194 TGGGCTTTCCGCGATGCAAGGCTATGCTCTGTCACCGGATGTGCGCATCGGTCGCGAG 1253
 QY 1368 CTGCGAGATTAATGAGATACAGATCTTGGAGTGGCTTACCAATTTGCCATGAGTCTG 1427
 Db 1254 CTGCACTGCAACCATGAGAGCGCTTCTCTCAGCGTTTGTGGTGGCCCATGAGACTG 1313
 QY 1428 ACACAACTTTGGCATGATTCATGATGAGAGGGAACATGTGTAAAGTTCGA----- 1481
 Db 1314 CCACGTGCTGGCATGAGACACGCGGACGAGGCAACCGCTGTGGGACGAGGTGGCGCT 1373
 QY 1482 GGGCAACATCANGTCCCTACATATGGCAGAGCGCAATGGAGTCTTCTCTGCTCACCGTG 1541
 Db 1374 GGGCAGCATCATGGCGCCCTGCTGGTGGAGCGCGCTTCCACCGCTTCCAGTGGTCCGCTG 1433
 QY 1542 CAGCGCGCATGATCTACACAAATTTCTTAAGCACCGCTCAAGCTATCTGCTTGTCTGATCA 1601
 Db 1434 CAGCCAGCAGAGGTGAGCGCTACCTGCACTCTTATGA-----CTGCTCTGATGA 1487
 QY 1602 GCCAAAGCCTGTGAAGAAATACAGATATCTGAGAAATTTGCCAGGAGAAATATATGATGAC 1661
 Db 1488 CCGCTTTGGCC---ACGACTGGCGGGCTGGCCCGAGCTCCGCGGACTGCACCTACTCCAT 1544
 QY 1662 AAACACACAGTGGGAGTGGCGATGTCGAGAGAGAAACCAAGCTCTGCACTGCTGACTTTAA 1721
 Db 1545 GAACAGCAATCGCTTTGACTTTCGCGCTGGGCTTACATGATGTGCAAGCGCTTCGGAC 1604
 QY 1722 AAAGCAATCTTAAGCCCTGTGTGCCATCTGATTTGGAAG-----AAATGTGAGAC 1775
 Db 1605 CTTTGACCTGCAAGCAGCTGTGTGCGACCATCTCTGCAACCCCTACTTTTTCAGAC 1664
 QY 1776 TAAATTTATGCCAGCAGAGAGGACAAATTTGTGGCATGACATGTGTGGCGGAGG 1835
 Db 1665 CAAGAAGGCGCCCTTGGAGCGGATATGTGTGCACTTGGCAAGCATTTTAAAG 1724
 QY 1836 ACAGTGTGTAATATGATGTAAGAGCCCGCAAGCCCAAGCCATGCGCATGTGCGACTG 1895
 Db 1725 ACATGCACTTGGTGCACCTGACATCTCAACCGGAG---GGCAGTGGGGCGCTG 1781
 QY 1896 GTCTTTGTGCTCCAGCATCTGCGAGGCGGAGTATCTCTAGAGAGTGCCT 1955
 Db 1782 GAGTCCGTTTGGCTCTGCTCACTGCTGCGAGCGGCGGTGAAGTTGAGGACCGCCCA 1841
 QY 1956 CTGCAACCCCGAGCATCTGATGAGAGAGTGTGTGAGGCTTCCACTCGCATCT 2015
 Db 1842 GTGTGACACCCACCGGCAAGCGGCGCCGACCTGCTGGGCTTGGCTAGACTT 1901
 QY 2016 GAAGCTGTGACAGTGAAGATTCGCGGACAGTGTGAGTTCGGTCTGCTGAGTG 2075
 Db 1902 CAGCTCTGAGCGCCAGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1961
 QY 2076 TGCBA---CCACACAGCAGAGATTTGAGAGGGGCGCATACAGAGTGCATACAC 2132
 Db 1962 CGGCTGTGGGACTGTACTTGTGACACCGGCGGCGCCACACACCTGCTGCTGCTGCTGCTG 2021
 QY 2133 TCAAGTAGAGATCAGACTTATGCAACTCTACTGTATGTCAGAGAGATTTGATTTCTT 2192
 Db 2022 GCACGGGATGCCAGAGAGATGTCACCTGTACTGTGCTGCTGCTGCTGCTGCTGCTGCTG 2081
 QY 2193 CTTTCTTTCTGCAATTAAGTCAAGAGTGGGACTCCATGCTC---GGAGGATAGCGTAA 2249
 Db 2082 GGTGTCCATGAAGCGCATGCTGATGATGATGAGAGCGGCTGCTGCTGCTGCTGCTGCTGCTG 2141
 QY 2250 TGTGTATAGATGGGATGATGAGAGATTTGATGTGCAATGTCTCTGATCTGATGCTG 2309
 Db 2142 CTTCTGTGCGCGGAGCTGACAGAGTGGGCTGTGACGCTGTGATGCTGCTGCTGCTGCTG 2201
 QY 2310 TGTGAAGAGCTGTGTGGGCTGTAAAGCGGAATTAAGTCTGAGCTGCGAGATTCAGAGGG 2369
 Db 2202 GCAGGAGACAAAGTGTGGCGTGTGGGAGGGGACACAGCCACTGCAAGTGGTCAAGGG 2261
 QY 2370 TCTTACACCAAGCACCAACCAACCAACCAAGTATTAATGATGCTGCTGCTGCTGCTGCTG 2429

RESULT 10

US-09-491-522-9
 ; Sequence 9, Application US/09491522
 ; Patent No. 6428998
 ; GENERAL INFORMATION:
 ; APPLICANT: Colige, Alain
 ; APPLICANT: Lepiere, Charles M.
 ; APPLICANT: Prockop, Darwin J.
 ; TITLE OF INVENTION: RECOMBINANT N-PROTEINASE,
 ; TITLE OF INVENTION: AND THE PRODUCTION, METHODS AND USES THEREOF
 ; NUMBER OF SEQUENCES: 17
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds, LLP
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: NY
 ; COUNTRY: USA
 ; ZIP: 10036-2811
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: Windows
 ; SOFTWARE: FastSeq for Windows Version 2.0b
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/491,522
 ; FILING DATE:

CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/886,333
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Abrams, Samuel B
 ; REGISTRATION NUMBER: 30,605
 ; REFERENCE/DOCKET NUMBER: 8389-0060-999
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 650-493-4935
 ; TELEFAX: 650-493-5556
 ; TELEX: 66147 PSNIE
 ; INFORMATION FOR SEQ ID NO: 9:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2489 bases
 ; TYPE: nucleic acid
 ; STRANDNESS: single
 ; TOPOLOGY: linear

US-09-491-522-9

Query Match 3.9%; Score 114.2; DB 4; Length 2450;

Best Local Similarity 47.8%; Pred No. 1.5e-22;

Matches 570; Conservative 0; Mismatches 593; Indels 30; Gaps 7;

QY 1309 GAGCCCTGTGACACTTTGGGATTTGACCCATTAAGTGGATGTGTAAATATGCGAGC 1368
 Db 1233 GGGCCCTTGGGATGCAAGGCTATGCTCTGTCATCTGGGATGTGCCACCCCGTCCGAGC 1292
 QY 1369 TGCAGATTAATGAAGATACAGAGTCTGGAGCTGGGCTTCCACCATTTGCCCATGAGTCTGGA 1428
 Db 1293 TGCACGCTGAACACAGAGAGCGCTTCTCTCTGCTTCTGCTGGTGGGCGCCACGAGACTGC 1352
 QY 1429 CACACCTTTGGCATGATTCATGATGAGAGGGAACATGTGTAAAGTTCGA-----G 1482
 Db 1353 CATGTGCTGGCATGAGCATGATGGCAGGCGCAACCGCTGCGGTGACGAGGTGGCGCTG 1412
 QY 1483 GGCACATCATGTCCCTACATTTGCGAGAGCGCAATGAGTCTTCTCTGCTGCTCAGCTGC 1542
 Db 1413 GGCAGCATCATGGCCCTCGTGTGCGAGGCGCTTCCATCGCTTCCATCGTCCGCTGC 1472

Db	1644	TTGACCCGTGCAAAACAGCTGTGGTGCAGCCACCTGACACACCCTACTTTTGCAGACA	1704
Qy	1777	AAA'TTTATCCGACGACGACAGGACACAA'TTTGTGGSCATGACATGTGGTGCAGGAGGA	1836
Db	1704	AGAAGGGGACACCCCTGTGATGGACCATGTGTGGCTTCGGCAGCACTGCTTTAAGGA	1763
Qy	1837	CAGTGTGTGAATATATGTTGATGAGGGCCCAAGGCCACCACTATGGCCACTTGGTCCGACTGG	1896
Db	1764	CAC'TGCATCTGGCTGCACACTGCACATTC'CAAC--GAGATGCCAACTGGGGTGCCTGG	1820
Qy	1897	TC'TTCTTGGTCCCATGCTCTCAGAGCACTCGGAGGGGGAGTAGTCTCATAGAGAGTGCCTC	1956
Db	1821	AGTCCCTTCGGCTCTCTCGCTACCTCGGSCACAGGTGTGAAGTTCAGCAACCGTCAG	1880
Qy	1957	TGCAACAACCCCAAGCCATCGCATGAGAGGGAGTTC'TGTAGGGCTCCCACTCGCACTCTG	2016
Db	1881	TGCGACAACCCCAACACCCAGCCAA'TGGGGGCCCACTATGCTCGGGCCTTCGCTTCAGT	1940
Qy	2017	TAGCTCTGCAACAGCTCAGAAATGTCCCCGGGACAGTGT'TGACTTCGGTCTGCTCA'GTG	2076
Db	1941	CAGCTCTGCAACTCGCAGAGACTGTCCTGTGACGGCTGGCCGCACTCCGCGAGAGCA'GTGC	2000
Qy	2077	GCCGA---GCACAACAGCAGACGATTCAGAGGGGGCAGCTACAA'GTGGGAAGCCTTACCT	2133
Db	2001	GGGACAGTGGGACCTGT'TACTTCGAGCATGTGTGACGCT'CAACACCACTGGTGC'CCCAAG	2060
Qy	2134	CAAGTAGAGATCAGGACTTATGC'AAACTCTACTGTATGCGAAGGATTTGA'TTCTTC	2193
Db	2061	CACCGGACGCCAAGGACGGTGTCTCTCTACTGTGAGTCC'CAAGAGACCGGGGAGGTG	2120
Qy	2194	TTTCTTTGTCAAA'TAAGTCAAAGATGGCACTCATG--CTGGAGGATAGCCGTAA'T	2250
Db	2121	GTGTCCATAGCGATATGTGTGCTGTGAGGACAGCTGTCTTCTACAGAGACGCTTC'GAC	2180
Qy	2251	GTTTGTATAGATGGATATGTGAGAGAGTGTGATGATGATCTTCTTGGATCTGATGCT	2310
Db	2181	CTCTGCTGGCTGGGAGCTGCAGAGAGATGGTGGCTGTGA'CGGGTGTCTGGCTTCAGCA	2240
Qy	2311	GTTCGAGAGCTCTGTGGGTGTGAACGGGATTAACCTCAGCTTCGACGATTAACAGGGGT	2370
Db	2241	CAGAGGACAA'GTGTGGTGTGTGTGGAGGGGAACTCCATCCATGC'GANGTGTCTAGGGC	2300
Qy	2371	CTCTACACCAAGCACACACACCAACAGTATAT'ACATGATTCACATGGTCA'CCATCTCTCGGA	2430
Db	2301	ACGTTCTCGCGTCGCCAAGAAGCTGGTGTATATCAAGTATTTGAGATCCCGCAGGA	2360
Qy	2431	GCCCGGAGTATCCGCTATCATGAATGAACGTCTCTACCTCTCATCTTCTGT	2483
Db	2361	GCCAGACA'CTCTCTAATTCAGAGAGACAGACCA'CCAGCACTCACTCTGGCGGT	2413

RESULT 13

```

RESOL: 12
US-09-369-364A-12
; Sequence 12, Application US/09369364A
; Patent No. 6391610
; GENERAL INFORMATION:
; APPLICANT: Apre, Suncel
; APPLICANT: Hurskainen, Tiina L.
; APPLICANT: Hirschata, Satoshi
; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
; FILE REFERENCE: 26473/4007/10-30-00
; CURRENT APPLICATION NUMBER: US/09/369,364A
; CURRENT FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 5804

```

Query Match	3.9%	Score 113.6	DB 4	Length 5804
Best local similarity	47.1%	Pred. No. 3.8e-22		
Matches	542	Conservative	0	Mismatches 561; Indels 33; Gaps 6;
QY	1245	AGATGGAGCTGTGTCAGACACAGCCGCTACTACTGCTCTGGATATATGTTCTCGATATATGTTCTCGGAA	1304	
DB	974	AGGTGGAGTTCATCATGATCTGCTGTTCTTCTTAACAGACAGATATCTG---CAGAC	1030	
QY	1305	GAATGAGCCCTGTGACACTTTTGGATTTGACCCATAGTGGATGTGTATGTAATATCG	1364	
DB	1031	TCAGCAAAATGTGATACCTTTAGGCTTGCTGAACTGGGAACCAITTTGATCCCTATAG	1090	
QY	1365	CAGCTGGAGATTAATGAAGATACAGGCTTTGGACTGGCTTTCACATTGGCCGATGCTC	1424	
DB	1091	AAAGCTGTTCTATTAGTGAAAGATAGTGATGATGATACMGCTTTTACGATCCCCATGCT	1150	
QY	1425	TGGAACAACTTTGGGCGTATCATGATGAGAGAGGAACATGTGTAAAGTCCGAGG	1484	
DB	1151	GGGCCATGTGTTAAACATGCTCATGATGACACAAATGTAAAGAGAGGATTA	1210	
QY	1485	CA-----ACATCATGTCCCTACATTTGGAGAGCGCAATGGAGCTTCTCTCTGGTCTC	1535	
DB	1211	GAGTCCCGACGATGATGGCTGCACACTGAACTTCTACACCAACCCCTGGATGTGGTCTC	1270	
QY	1536	ACCTCGAGCGCCAGTATCTACACAAATTTCTTAAGACCGCTCAGAGCTATCTGCTTCG	1595	
DB	1271	AAAGTGTAGTGAAAATATATCATCTGAGTGTATTAAGACATGAGTTATGGGAGTGTTCCT	1330	
QY	1596	TGATCAGCCAAAGCCTGTGAGGAATATCAAGTATCTCTGAAATATGCGACAGGAATATATA	1655	
DB	1331	TAAAGAACCTGATCT---CAGACCTTACCCTTTGCTGTGCACATGCCAGGACCTCTTTTA	1387	
QY	1656	TGATGCAACACACAGTGTCAAGTGGCAGTTCGGAGAGAAAGCAAGCTCTGCATGCTGGA	1715	
DB	1388	CACGCTGAAATAACAATGNGAATATGATTTTGGACAGAGTCTCAGGTGTGCCCATATAT	1447	
QY	1716	CTTTAAAAAGGACATCTGTAAGGCCCTGTGGTGCCATCGTATTTGGAAGGAATATGGAC	1775	
DB	1448	GATCGAGTCAGACGGCTCTGGTGCAATAACGTCATATGGAGTACACAAAGCTCCCGAC	1507	
QY	1776	TAAATTTATGCGACGACAGAGGCAACATTTGTGGGCATGACATGTGTGTGCGGGGAG	1835	
DB	1508	TCAGCACACACCTTGGGCGCATATGGAGCGAGTGGAGCCTTGGAAAGCACTCTCAAGNATGG	1567	
QY	1836	ACAGTGTGTAAATATGTGTATGAAGGCCCAAGCCACCACTGGCCACTGGTTCGGACGTG	1895	
DB	1568	ATTTTGTGTCTCCAAAGAAATGGATGTGCCGTG---ACAGATGGATCTCTGGGGAAGTGT	1624	
QY	1896	GTCTTCTTGTTCCCATGCTCCAGGACCTCGCGAGGGGGAGTATCTCATAGGAGTCGGCT	1955	
DB	1625	GAGTCCCTTTGGAAACCTGTCTCAGACATGTGTGGAGGGGSCATCAAAACAGCCATTCGAGA	1684	
QY	1956	CTGCACCAACCCCAAGCCCATGCGATGGAGGGAAGTTTGTGTGAGGGCTTCCATCTGCACACT	2015	
DB	1685	GTGCAACAGACCCAGAACCAAAAAATGGTGGAAAAATAGTGTGTAGGACGTAGAAATGAAT	1744	
QY	2016	GAGGCTCTGCAACAGTCAGAAATGTCCCGGGGACAGTGTGTACTTCCGTGCTGCTCAGTG	2075	
DB	1745	TAACTCTCTCAACACCGGAGCCATGTCTCAAGCAGAGACGAGACTTCCGAGATGAACAGTG	1804	
QY	2076	TGCCGAGCACACAGCAGACAGT-----TCNAGAGGGCGGCATCTACAGTGGGAAGCT--	2127	
DB	1805	TGCTCACTTTTGAACGGGAAGCATTTTAAACATCAACGGTGTGTTCTCCCATATGTGCGTGGT	1864	
QY	2128	-----TACACTCAAGTAGAAGATCAGAGCTTATGCAAACTCTACTGTATCGGCAGAGG	2180	

```

Db 1865 CCTAATACAGTGGATTCCTGATGAGGACCGGTGCAAGTTGTTCTGCAGAGTGGCAGG 1924
QY 2181 ATTGATTTCTCTTTTCTTTCTCAATAAAGTCAAGATGGGACTCCATCTCGGAGGA 2240
Db 1925 GAACACAGCCTACTACTCAGCTTCGAGACAGAGTATAGATGGAACCTCTTGTGGCCAGGA 1984
QY 2241 TAGCCGTATGTTTGTATAGATGGGATATGAGATGAGTGGATGTGACAAATGTCCTTGG 2300
Db 1985 CACAAATGATATCTGTCTCAGGCGCTTTCGCGGCAAGCTGGATGCGATCATGTTTAA 2044
QY 2301 ATTCGATGCTGTGAGAGAGCTGTGGGGTGTGTAAACGGCAATTAATCTCAGCCTGCA 2356
Db 2045 CTCAAAAGCCCGAGAGATAAATGCGGGGTTTGTGGCGGATTAATTTCTTATGCA 2100

RESULT 13
US-09-369-364A-1
; Sequence 1, Application US/09369364A
; Patent No. 6391610
; GENERAL INFORMATION:
; APPLICANT: Apte, Suneel
; APPLICANT: Hurskainen, Tiina L.
; APPLICANT: Hirohata, Satoshi
; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
; FILE REFERENCE: 26473/4007/10-30-00
; CURRENT APPLICATION NUMBER: US/09/369,364A
; CURRENT FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1
; LENGTH: 3002
; TYPE: DNA
; ORGANISM: mus musculus ADAMTS-5
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (18)..(2810)
US-09-369-364A-1

Query Match 3.5%; Score 102.4; DB 4; Length 3002;
Best local similarity 48.0%; Freq. No. 4.3e-15;
Matches 538; Conservative 0; Mismatches 541; Indels 41; Gaps 7;

QY 1296 TTCCTGGAGATAGAGCCCTGTGACACTTTGGATTTGGACCATTAAGTGGGAATGTAG 1355
Db 1124 TTATGTGGGCATCTTCAATGTGACGCTTGGGATGGCAGAGCTTGGACCATATATGTC 1183
QY 1356 TAAATATCGCAGCTGACGATTAATGAAGTACAGGCTTGGAGCTGGCCTTCACCATTCG 1415
Db 1184 TCGGAGGCGGAGCTTGGAGTGTGAGATGATGAGTGGCTCCATGACGCTTCACTGTGGC 1243
QY 1416 CAGTGTGAGAGCACTTTGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1463
Db 1244 TCTGAAATTTGGGATCTACTTGGGCTTCTTCAATGAGTTCCTGAGAGAA 1303
QY 1464 CATGTGTAAAGATTCGAGGGGCAATCATATGTCCTTACATTTGCGAGAGCGAATGGAT 1523
Db 1304 CTTTCGTACTACAGAGCAAGCGTTTAAATGCTTCAATCTTACAGCATCGATGCATC 1363
QY 1524 CTTCTCTGTGTACCTGCGAGCGGCTAGTATCTACAGAAATTTCTAAGCAGCGCTCAAGC 1583
Db 1364 CAAAGCCCTGGTCTCAATGCGAGCTGAGGAGGAGTTCAGAGATTTCTGGATGATGGTCATGG 1423
QY 1594 TATCTGCTGTCTGATGAGCAAGCTGTGAGGAGTACAGATATCTAAGCAAAATTTGCC 1643
Db 1424 TAATGTTTCTAGACCTTACCAG-----GAAGCAGATTTTGGGTCCCGAGAACTCCC 1477
QY 1644 AGGAGATTTATGATGCAACACACAGTGCATGCGAGTTCGGAGAGAAAGCAAGC 1702
Db 1478 AGGACAGACCTTACGATGCCACACAGCAGTGCATGCACTTGCATTTGGGCTCGATCTCGGT 1537
QY 1703 TCTGATCTCGGACTTTAAAGAGGACATCTGTAAAGCCCTGCTGGTGCATCTGATTGAA 1762
Db 1538 GTCCCTTGGCATGATGTCTGTGCGGGCTGTGGTGTCTGTGTGTGCTGCGGAGCCAAA 1596

```

```

QY 1763 GGAAATGTGAGACTAAATTTATTCAGCAGCAGCAAGGCAACAAATTTGTGGCATGACATGT 1822
Db 1597 TGGTGTGTCTGACCAAGAACTGCGGCTGTGGAGGCACTCCCTGTGGGAGGAAGAG 1656
QY 1823 GTGCGCGGGGAGGACAGTGTGTGAAATATGATGAA-----GGCCCCAGAGCCACCC 1876
Db 1657 TCTGCTTTCAGGCGAAATGTGTGCAAAACTAAGAAAAATATTTACTTGACATCAGGCC 1716
QY 1877 ATGGCCACTGGTGGAGCTGTCTTCTTGTGGTCCCATGCTCCAGACCTTCGAGAGGGGAG 1936
Db 1717 ATGGAAATTTGGGGGTCTGGGGCCCTGGGGTCTGAGTGTCTGCTCATGCGGGGAGGAG 1776
QY 1937 TATCTCAAGAGTGGGCTCTGACCAACCCCAAGCCCATGCAATGAGAGGAGTTCGTGTG 1996
Db 1777 TGCAGTTTGGCTTACCGCATTTAAATACCTTGCACCTCGCAAAAGTGGCGCTACTGCA 1836
QY 1997 AGGGCTCCACTCGCACTCTGAGCTCTGCAACGTGCAAAATGTCCCGGAGCACTGTG 2056
Db 1837 CAGGAGAGGGGCCATATACCGTCTTCTGAGTGTTCACCTGCCCAACCCATGTTAAAT 1896
QY 2057 ACTCGGTGTCTGTGAGTGTGGGAGCAGCAACAGCAGAC-----GATTGAGAGCGGCA 2111
Db 1897 CTTTGGCCATGAGCGGTGTGAAGCCAAATGCTATCATGCTGTGCAAAAGAGTCA 1956
QY 2112 CTCAAGTGGAGCCCTTACCTCAAGTGAAGATC-----AGCACTTATGCAAC 2161
Db 1957 AACAATTTGTAGATGGGTTCCTCCCAATATGCAAGTGTCTGCGGAGATGTGTCAAGC 2016
QY 2162 TCTACTGTATCGCAGAGATTTGATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2221
Db 2017 TTACTTCAGAGCTAAGGCGCAGGCTACTGTGTGCTTCTTCTTCTTCTTCTTCTTCTTCT 2076
QY 2222 GGACTCTGCTGCGAGTATGCGGTAAATGTTGTATAGATGGATGTCTCAGAGTGTG 2281
Db 2077 GGAATGATGAGCGCGCTGAGCACTCTGTGTGCTGCGAGGAGCGTGTGTGAGACTG 2136
QY 2282 GATGAGCAATGCTCTTGGATCTGATGCTGTGAGAGCTGTCTGTGTGTGTGTGTGTGTGT 2341
Db 2137 GATGTGCGCATTTGCTCAAGCTACAAATATGACAGTGTGGAGTGTGGGAGGGG 2196
QY 2342 ATACTTCAGCTCAGGATTCACAGGCGTCTCTACACCA 2381
Db 2197 ATACTTCAGTGTACAGAGTATTCGAGACCTTCAATAA 2236

```

```

RESULT 14
US-09-445-023A-13
; Sequence 13, Application US/09445023A
; Patent No. 656568
; GENERAL INFORMATION:
; APPLICANT: Hirose, Kunitaka
; APPLICANT: Inoguchi, Biji
; APPLICANT: Hakezaki, Michinori
; APPLICANT: Ishioke, Keiko
; APPLICANT: Ishida, Yukako
; APPLICANT: Matsushima, Kouji
; APPLICANT: Kuno, Kouji
; TITLE OF INVENTION: Human ADAMTS-1 protein, gene encoding the same, pharmaceutical
; FILE REFERENCE: Q57092
; CURRENT APPLICATION NUMBER: US/09/445,023A
; CURRENT FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: JP 9-160422
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent In version 3.0
; SEQ ID NO 13
; LENGTH: 2184
; TYPE: DNA
; ORGANISM: Mus sp.
; FEATURE:
; NAME/KEY: exon

```

; LOCATION: (1)..(2184)
US-09-445-023A-13

Query Match 3.5%; Score 101.2; DB 4; Length 2184;
Best Local Similarity 48.3%; Pred. No. 86-19; Indels 51; Gaps 8;
Matches 564; Conservative 0; Mismatches 553;

QY 1313 CTGTGACCTTGGATTTGCAACCAATAGTGGATGTAATATCCAGCTGCA 1372
DB 377 CGTGTGACCTTGGATTTGCAACCAATAGTGGATGTAATATCCAGCTGCA 436
QY 1373 CGATTTAAGATGATGCTTCCACCAATAGTGGATGTAATATCCAGCTGCA 1432
DB 437 CAGTCTAAGATGATGCTTCCACCAATAGTGGATGTAATATCCAGCTGCA 496
QY 1433 ACTTTGCAATGATCATGATGAGAGGAGAC-----ATGTGTAAAGTGG 1480
DB 497 TGTTTAAGTGGCGCAGATGATGTAAGCACTGTGCCAGCTTGAATGTGTGAGTGGCG 556
QY 1481 AGGCGACATCATGTCCTCCATACATTTGGCAGGACGCAATGAGTCTTCTCTGGTCACTCT 1540
DB 557 ATTCTCATCTGATGCGCTCGATGCTCTCCAGCTTAGACCATAGCAGCCCTGGTCACTCT 616
QY 1541 GCAGCGCGCATGATCTACACAAATTTCTAAGCACCGCTCAAGCTATCTGCTTGTCTGATC 1600
DB 617 GCAGTGGCTACATGTCACGTCTCTCTCTAGATAATGACACAGGGAATGTTTGTATGACA 676
QY 1601 AGCCAAAGCCTGTGAAGGATACAAATGATCTCTGAGAAATTTGCCAGAGATTTATATGATG 1660
DB 677 AGCCCCA-----GAATCCAATCAAGCTCTCTGATCTTCCCGGTACTCTGTACGATG 730
QY 1661 CAAACACACAGTGCAGAGTGGCAGTTCGGAGAGAAAGCAAGCTCTGCAATGCTGG--ACTT 1718
DB 731 CCAACCGCCAGTGTCAATTTACATTTGGAGAGGAATCCAGACCTGCGCTGTATGACGCA 790
QY 1719 TAAAGAGACATCTGAAAGCCTGTGTGTCCATC-GTATTTGAAAGAAATCTGAGACTA 1777
DB 791 GCACATGATCTACCTCTGTGTGCACTGGCAGCTCCGGTGTCTTACTGTGTCCCAACAA 850
QY 1778 AATTTATGTCAGACAGAGAGGACAAATTTGTGGCATGACATGTGTGGCGCGGAGAGAC 1837
DB 851 AACCTTCCCTTGGCGAGATGGCAGCCAGCTGTGGAGAGGAGGTGTGTGTGTCAGTGGCA 910
QY 1838 AGTGTGTGAA-----ATTTGTGTATGAGGCGCCCAAGCCACCCATGGCCATGTGTGCG 1891
DB 911 AGTGTGTGAAAG 970
QY 1892 ACTGTGTCTTGTGTGTCCTCCAGAGCTCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1951
DB 971 CATGGGAG 1030
QY 1952 GCCTCTGACCAACCCCAAGCAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2011
DB 1031 GAGATGTGACAAACCCAGTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1090
QY 2012 CTCTCAGCTCTGCAAGTCAAGATGTGTC-----CGGAGAGTGTGATCTTCCG 2063
DB 1091 GCTTACAGTCTCTTAACTAG 1150
QY 2064 TGTCTCTCAGTGTGGCAGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2123
DB 1151 AGCAGTGTGGAGCGCAGCAATGAGTCTTCCAAAGCTCTCTTGGAGATGAGCCCACTGTAG 1210
QY 2124 GCCTTACACTCAAGTGAAGATC-----AGGACTTTATGCAAACTCTACTGTATGCG 2173
DB 1211 AGTGGACACCAAGTACCGCGCGCTGTGCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1270
QY 2174 CAGAGAGATTTGATTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 2233
DB 1271 CCAAGAGCATGCTGCTATCTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 1330
QY 2234 CGAGAGATAGCCGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 2293

DB 1331 GTCCAGACTTACCTTGTCTGTGTGCAAGGAGCTGTGTGAAGCTGGCTGTGATCGCA 1390
QY 2294 TCTTTGATCTGTGTGTGGAAGAGCTGTGTGGGTGTGTAAACGGGAATTAATCACTCAGCCT 2353
DB 1391 TCTAGACTTCAAAAGAGAGTGTGTAAAGTGTGGCGTTTGTGGAGAGAAAGGTTCACAT 1450
QY 2354 GCAGCATTTACAGGGGCTTCTACACAGGACCCACACCAACCATGATTTATCATATGG 2413
DB 1451 GCAGAGAGATGTGAGGAATAGTCACTA-----GTACAGAGACTGGGTATCATGATCTG 1504
QY 2414 TCACCATCTCTCTGAGCCCGGGGAT 2441
DB 1505 TCACATCTCTCTGAGCCACCAACAT 1532

RESULT 15

US-09-484-970B-58
; Sequence 58, Application US/09484970B
; Patent No. 6426186
; GENERAL INFORMATION:
; APPLICANT: Jones, Karen A.
; APPLICANT: Volkmut, Wayne
; TITLE OF INVENTION: BONE REMODELING GENES
; FILE REFERENCE: PB-0014 US
; CURRENT APPLICATION NUMBER: US/09/484.970B
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: PEEL Program
; SEQ ID NO 58
; LENGTH: 3706
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6426186 007074.1
; NAME/KEY: unsure
; LOCATION: 3634, 3638-3639, 3642-3643, 3647-3648, 3652, 3654-3658, 3664, 3674, 3681
; OTHER INFORMATION: a, t, c, g, or other
US-09-484-970B-58

Query Match 3.4%; Score 97.2; DB 4; Length 3706;

Best Local Similarity 48.2%; Pred. No. 1.5e-17;
Matches 523; Conservative 0; Mismatches 518; Indels 45; Gaps 7;

QY 1310 AGCCCTGTGACCTTTGGGATTTGCCACCATAGTGGATGTAGTAAATATCCAGCT 1369
DB 526 AGACATGTGATCTCTGGGATGCTGATGTGGAACTGTGTGTGATCCAGCAGAGCT 585
QY 1370 GCAGCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1429
DB 586 GCTCGCTAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 645
QY 1430 ACACATTTGGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1477
DB 646 ACCTGTTACATGCTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 705
QY 1478 CGAGGCGCAACATCATGTCCTTACATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1537
DB 706 AGGATTTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 765
QY 1538 CTTGAGCGCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1597
DB 766 CTTGAGCGCTTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 825
QY 1598 ATGACCAAAAGCTGTGAGAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1657
DB 826 A-----CAGCTCAGAGATCCATACAGCTCCAGCGATCTCTCTGGCACTCGTACG 879
QY 1658 ATGCAAAACAGAGTGTGAAGTGTGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1714
DB 880 ATGCCCAACCGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 939

QY 1715 ACTTTAAAGGACATCTGTAAAGCCCTGTGGTCCATGTTATGGAGCAATGTGAG 1774
Db |||||
940 CCAGCACATGTAGCACCTTGTGTGTACCGCACCTCTGTGGGTCTGTGTGTCTAAA 999
QY 1775 CTAAATTTTGCAGCAGCAGAGCACAATTTGTGGGCATGACATGTGTGCGGGGAG 1834
Db |||||
1000 CCAAACACTTCCCGTGGCGGATGGCACACAGCTGTGAGAGGGAAATGGTGTATCAAG 1059
QY 1835 GACAGTGTGTGAATATGTGTATGAGGCC- - - - -AAGCCACCCATGGCCACTGT 1888
Db |||||
1060 GCAAGTGTGTGAACAAACCGACAGAAACATTTTGATACGCTTTTCATGGAGCTGG 1119
QY 1889 CGGACTGTGCTTTCTTGTCTCCCATGCTCCAGGACCTCGGAGGGGAGTATCTCATAGGA 1948
Db |||||
1120 GAATGTGGGGCCCTTGGGAGACTGTTCCGAGAACGTCCGCTGGAGGATCCAGTACAGA 1179
QY 1949 GTCCCTCTGCACCAACCCCAAGCCATCGCATGGAGGGAAGTTCTGTGAGGGCTCCACTC 2008
Db |||||
1180 TGAGGGAATGTGACAAACCCAGTCCCAAGAATGGAGGAAGTACTGTGAAGGCAAGAG 1239
QY 2009 GCACTCTGAAGCTCTGCAACAGTCAGAAATGTCCCGGACAGTGTGA- - -CTTCCGTG 2065
Db |||||
1240 TGCCTACAGATCTCTGTAACTTGAGGACTGTCCAGACAAATATGGAANAACCTTTAGAG 1299
QY 2066 CTGCTCAGTGTGCGGAGCACAA- - - - -CAGCAGACGATTCAGAGGGGCGGCACTACAA 2117
Db |||||
1300 AGGAACAATGTGAAGCACACACAGAGTTTTCAAAAGCTTCTTTGGAGTGGGCTGCGG 1359
QY 2118 GTGGAAGCT- - - - -TACACTCAGTAGAAGATCAGGACTTATGCBAATCAAGATGGACTCCAT 2170
Db |||||
1360 TGGAAATGGATTCCCAAGTACGTGGCGTCTCACCANAAGCACAGGTGCAAGCTCATCTGCC 1419
QY 2171 TCGCAGAGAGGATTTGATTTCTTCTTTTCATTGTGCAATAAGTCAAGATGGACTCCAT 2230
Db |||||
1420 AAGCCAAAGGCATTTGGCTACTTCTTGTGTTGACGCCCAAGGTGTAGTGGTACTCCAT 1479
QY 2231 GCTCGGAGTAGCGGTAAATGTTGTATAGTGGGATATGTGAGAGAGTTGGATGTGCA 2290
Db |||||
1480 GTAGCCAGATTCCACCTCTGTCTGTGTGTCAGAGACAGTGTGTAAAGCTGTGTGTGTC 1539
QY 2291 ATGTCCTTGGATCTGTGCTGTGAAGAGCTGTGTGGGGTGTGTAAAGGATTAAGTCTAG 2350
Db |||||
1540 GCATCATAGACTCCAAAAGAGTTTGATTAATGTGTGTGTTGCGGGGGAATGATCTTA 1599
QY 2351 CCTGCA 2356
Db |||||
1600 CTGTGA 1605

Search completed: October 28, 2003, 23:26:41

Job time : 208 secs

Result	No.	Score	Query Match	DB Length	ID	Description
	1	1874.6	64.8	3675	13	US-10-217-774-3 Sequence 3, Appl
	2	1874.6	64.8	3675	15	US-10-296-616-1 Sequence 1, Appl
	3	1874.6	64.8	4062	13	US-10-217-774-5 Sequence 5, Appl
	4	1874.6	64.8	3666	12	US-10-340-545A-1 Sequence 1, Appl
	5	673.4	23.3	3589	14	US-10-226-560-1 Sequence 1, Appl
	6	673.4	23.3	3013	14	US-10-226-560-3 Sequence 3, Appl
	7	663	22.9	1476	13	US-10-217-774-1 Sequence 1, Appl
C	8	292.8	10.1	525	12	US-10-029-386-8689 Sequence 8689, Ap
C	9	292.6	10.1	348	12	US-10-029-386-22417 Sequence 22417, A
	10	267.4	9.2	339	10	US-09-633-381-1437 Sequence 1, Appl
	11	265.8	9.2	2052	14	US-10-103-377C-1 Sequence 1, Appl
	12	265.8	9.2	3377	14	US-10-103-377C-3 Sequence 3, Appl
	13	241.2	8.3	3217	14	US-10-226-560-6 Sequence 6, Appl
	14	237	7.8	3274	10	US-09-563-731-23 Sequence 23, Appl
	15	227	7.8	3727	10	US-09-563-731-1 Sequence 1, Appl
	16	227	7.8	3879	10	US-09-788-043C-2 Sequence 2, Appl

APPLICANT: HILBUR, EYR
TITLE OF INVENTION: NO H520020393583A1e] Human proteases and polynucleotides encoding

QY 382 CTCACCTCAATATGACCTGGTCTCTGCTTACAGAGTTTGACCAACAGGGGGCGATTACGTGTCC 441
DB 176 -----AATATGAGCCTGGTCTCTGCTTACAGAGTTTGACCAACAGGGGGCGATTACGTGTCC 228
QY 442 CATGAATCATGCAACATCAGCGCGGAGAGACAGTGGCGGTGTCGAGGTTGAGTCT 501
DB 229 CATGAATCATGCAACATCAGCGCGGAGAGACAGTGGCGGTGTCGAGGTTGAGTCT 288
QY 502 CTTCACTTCGGCTGAAAGGCCCAAGCAACGACTTCCATCGATCTGAGGACTTCCAGC 561
DB 289 CTTCACTTCGGCTGAAAGGCCCAAGCAACGACTTCCATCGATCTGAGGACTTCCAGC 348
QY 562 AGCCTAGTGGCTCTGGCTTTATTTGTGAGACGTTGGGAAGACAGGCACTAAGTCTGTG 621
DB 349 AGCCTAGTGGCTCTGGCTTTATTTGTGAGACGTTGGGAAGACAGGCACTAAGTCTGTG 408
QY 632 CAGACTTTACCGCAGAGGACTTCTGTTTATCAAGGCTCTTTGGGATCAACAGAAAC 681
DB 409 CAGACTTTACCGCAGAGGACTTCTGTTTATCAAGGCTCTTTGGGATCAACAGAAAC 468
QY 682 TCGCMTGCGATGGAAGGAGTCTGTGAGGGTCCACTGCACTCTGAGACTTCCAAC 741
DB 469 TCCCTCA-----GTGGCCCTTTCAACTGCGCAAGGCTTGCA 504
QY 742 AGTCAGAAATGCTCCCGGACAGTGTGACTTCGCTGCTCAAGTGTGCGGAGCAAC 801
DB 505 GCATGATAGCAACAGAGAGGAGATGACTTCTTAAGGCCACTTCTTCAACCTTCA 564
QY 802 AGCAGCAATTTCAAGAGCGGCACTCAAGTGG--AAGCCTTACACTCAAGTAGAGCG 859
DB 565 TGGMACTCGGCGAGAGTGGCCAGGAGCTGCGCATCCACGTACTGTACAAGAGATCC 624
QY 860 ACTTATGCAACTCTACTGATGCGAGAGGATTTGATTTCTTCTTTCTTTGTCAAATA 919
DB 625 ACAGAGCCCATCTCTGGGGCCAGTAGAGTCTCTGGTGAACCTC-----AAGG 672
QY 920 AAGTCAAAGATGGAATCCATGCTCGAGGATAGCGTAATGTTGTATAGATGGGATAT 979
DB 673 ACATGGAGCTGGCACTCAACCTCCACAGCAGCGGACCTTGGCCCTGGGACTGCCCAA 732
QY 980 GTGAGCTCAGTGTGTGTCACATGTC--GCATGTCGCCAGCTCCCAAGGAAGACCTC 1038
DB 733 AAGCAGATTCTGTGGAAAGCCCAAGAAATACATGCCCCAGCTCCCAAGGAAGACCTC 792
QY 1039 TTCTATGCGCAGATGATTAAGTCTGTGTTACGGCATTAAGCGTCTCTCTTGAGGTCC 1098
DB 793 TTCTATGCGCAGATGATTAAGTCTGTGTTACGGCATTAAGCGTCTCTCTTGAGGTCC 852
QY 1099 CATGAATATGAGAACTCAAGCTGGAGACCTTGGTGGTGTGCGAAGAAAGATGATGCA 1158
DB 853 CATGAATATGAGAACTCAAGCTGGAGACCTTGGTGGTGTGCGAAGAAAGATGATGCA 912
QY 1159 AACCATGGCCATCAAAATATCAACCTTACGCTCCAGATACTCAATGTGTATCTGCT 1218
DB 913 AACCATGGCCATCAAAATATCAACCTTACGCTCCAGATACTCAATGTGTATCTGCT 972
QY 1219 TTATTTCAAGA----- 1229
DB 973 TTATTTCAAGATGAAACATAGGAGGAACATCAATGCAATGTGAGTCTGATCTT 1032
QY 1230 ----- 1229
DB 1033 CTAGAAGATGAACCGCGGACTGGTGTAGTAAAGTCAACAGCAGCAACATTAAGTAGC 1092
QY 1230 -----TGGATTGATGGGAAGATGGGACTCGTCAATGACCGCCGATC 1272
DB 1093 TTCTCCAGTGGCGATCTGATTTGATGGGAAGATGGGACTCGTCAATGACCGCCGATC 1152
QY 1273 TTACTGACTGGCTGATATGATCTGCTGCGAAGATGAGCGCTGTGACATTTGGGATTT 1332
DB 1153 TTACTGACTGGCTGATATGATTTCTTGGAGATGAGCCCTGTGTGACATTTGGGATTT 1212

QY 1333 GCACCCATAAAGTGGAAATGTGTAGTAAATATCGCAGCTGCACGATTAATTAAGATACAGGT 1392
DB 1213 GCACCCATAAAGTGGAAATGTGTAGTAAATATCGCAGCTGCACGATTAATTAAGATACAGGT 1272
QY 1393 CTTTGAAGCTGGCCCTTCAACATTTGCCATCGAGTCTGGGACACAACTTTTGGCATGATTCAATGAT 1452
DB 1273 CTTTGAAGCTGGCCCTTCAACATTTGCCATCGAGTCTGGGACACAACTTTTGGCATGATTCAATGAT 1332
QY 1453 GGAGAAAGGAAATGTGTAAAAAGTCCGAGGGCAACATCATGTGTCCCTTACATTTGGCAAGA 1512
DB 1333 GGAGAAAGGAAATGTGTAAAAAGTCCGAGGGCAACATCATGTGTCCCTTACATTTGGCAAGA 1392
QY 1513 CGCAATGGAGTCTTCTCTCGTCCACCCTTGAGCGCCGACAGTATCTACAACAAATTTCTTAAGC 1572
DB 1393 CGCAATGGAGTCTTCTCTCGTCCACCCTTGAGCGCCGACAGTATCTACAACAAATTTCTTAAGC 1452
QY 1573 ACCGCTCAAGCTATCTGCTGCTGATCAGCCAAAGACCTGTGAAGGAATACAAGTATCCT 1632
DB 1453 ACCGCTCAAGCTATCTGCTGCTGATCAGCCAAAGACCTGTGAAGGAATACAAGTATCCT 1512
QY 1633 GAGAAATTTGCCAGAGAAATATATGATGCAAAACACACAGTGCAGTGGCAGTTTGGAGAG 1692
DB 1513 GAGAAATTTGCCAGAGAAATATATGATGCAAAACACACAGTGCAGTGGCAGTTTGGAGAG 1572
QY 1693 AAAGCCAAAGCTCTGCATGCTGGACTTTTAAAGAGGACATCTGTAAAGCCCTGTGTGCCAT 1752
DB 1573 AAAGCCAAAGCTCTGCATGCTGGACTTTTAAAGAGGACATCTGTAAAGCCCTGTGTGCCAT 1632
QY 1753 CGTATTGGAAAGAAATGTGAGACTAAATTTTATGCCAGCAGCAGAAAGCACAATTTTGTGG 1812
DB 1633 CGTATTGGAAAGAAATGTGAGACTAAATTTTATGCCAGCAGCAGAAAGCACAATTTTGTGG 1692
QY 1813 CATGACATGTGTGTCGGGGGAGGACAGTGTGTAATATGTTGATGAAGGCCCAAGGCC 1872
DB 1693 CATGACATGTGTGTCGGGGGAGGACAGTGTGTAATATGTTGATGAAGGCCCAAGGCC 1752
QY 1873 ACCCATGGCCACTGTGTGGACTGTCTTCTTGTGTCCTCATGCTCCAGGACTTCCGGAGG 1932
DB 1753 ACCCATGGCCACTGTGTGGACTGTCTTCTTGTGTCCTCATGCTCCAGGACTTCCGGAGG 1812
QY 1933 GGAGTATCTCATAGGAGTGCCTCTGCAACCAACCCCAAGCCATCGCATGGAGGAAGTTTC 1992
DB 1813 GGAGTATCTCATAGGAGTGCCTCTGCAACCAACCCCAAGCCATCGCATGGAGGAAGTTTC 1872
QY 1993 TGTGAGGGCTTCCACTTCCACTCTGAAGCTCTGCAACAGTCAAGAAATGTCCCAGGACAGT 2052
DB 1873 TGTGAGGGCTTCCACTTCCACTCTGAAGCTCTGCAACAGTCAAGAAATGTCCCAGGACAGT 1932
QY 2053 GTTGACTTCCGTGTGTCTCAGTGTGCGGACCAACAGCAGACAGTTCAGAGGGCGGCAC 2112
DB 1933 GTTGACTTCCGTGTGTCTCAGTGTGCGGACCAACAGCAGACAGTTCAGAGGGCGGCAC 1992
QY 2113 TACAAGTGGAGGCTTACACTCAAGTGAAGATCAGGACTTATGCAAACTCTACTGTATC 2172
DB 1993 TACAAGTGGAGGCTTACACTCAAGTGAAGATCAGGACTTATGCAAACTCTACTGTATC 2052
QY 2173 GCAGAGGATTTGATTTCTTTCTTTTGTCTCAAATTAAGTCAAGAGATGGGACTCCATGC 2232
DB 2053 GCAGAGGATTTGATTTCTTTCTTTTGTCTCAAATTAAGTCAAGAGATGGGACTCCATGC 2112
QY 2233 TCGAGGATAGCCGTAATTTGTTATAGATGGGATATGTGAGAGAGTTGGATGTGCAAT 2292
DB 2113 TCGAGGATAGCCGTAATTTGTTATAGATGGGATATGTGAGAGAGTTGGATGTGCAAT 2172
QY 2293 GTCTTGGATCTGATGCTGTGGAAGAGCTGTGTGGGTGTGTAAAGGGAATTAATCTAGCC 2352
DB 2173 GTCTTGGATCTGATGCTGTGGAAGAGCTGTGTGGGTGTGTAAAGGGAATTAATCTAGCC 2232
QY 2353 TGCACGATTCACAGGGGTCTCTACA CCAAGACACCA CCAACCAAGTATTATCATG 2412
DB 2233 TGCACGATTCACAGGGGTCTCTACA CCAAGACACCA CCAACCAAGTATTATCATG 2292
QY 2413 GTCCACCATTTCTTCTGAGGCCGGAGTATCGCATCTATGAAATGAAGCTCTCTACTCTCC 2472

QY	1453	GGAGAGGGGACCATGCTCTAATAAGCTCCGAGGGGACCATCATCTCCCTCCCTACATTTCCGACGA	1513
DB	1333	GGAGAGGGGACCATGCTCTAATAAGCTCCGAGGGGACCATCATCTCCCTCCCTACATTTCCGACGA	1393
QY	1513	CGCAATCGGAGTCTTCTCTGCTGGTACCCCTCGAGCGCCACATATCTACACAAATTTCTTACGG	1573
DB	1393	CGCAATCGGAGTCTTCTCTGCTGGTACCCCTCGAGCGCCACATATCTACACAAATTTCTTACGG	1453
QY	1573	ACCGCTCCAGAGTATCTGCCTCTGCTGATCAGCCCAAGCCCTGTGAGAGGATACAGATATCTCT	1633
DB	1453	ACCGCTCCAGAGTATCTGCCTCTGCTGATCAGCCCAAGCCCTGTGAGAGGATACAGATATCTCT	1513
QY	1633	GAGAAATTCGACGAGAGAAATTAATGATGCAAAACACACAGCTGCAGAGGATACAGATATCTCT	1693
DB	1513	GAGAAATTCGACGAGAGAAATTAATGATGCAAAACACACAGCTGCAGAGGATACAGATATCTCT	1573
QY	1693	AAAGCCCAAGCTCTGCATGCTCGAGCTTTTAAAGAGGACATCTGTAAAGCCCTGTGTGCGCAT	1753
DB	1573	AAAGCCCAAGCTCTGCATGCTCGAGCTTTTAAAGAGGACATCTGTAAAGCCCTGTGTGCGCAT	1633
QY	1753	CGTATTCGAGAGCAATCGAGACTAATATTATGCGACGACGAGAGGCGCAATTTGTGGG	1813
DB	1633	CGTATTCGAGAGCAATCGAGACTAATATTATGCGACGACGAGAGGCGCAATTTGTGGG	1693
QY	1813	CATCAGCATGTGCTCGCGGAGGAGCATGTGTGTGAATATGCTGATGAAGGCCCCACAGCCC	1873
DB	1693	CATCAGCATGTGCTCGCGGAGGAGCATGTGTGTGAATATGCTGATGAAGGCCCCACAGCCC	1753
QY	1873	ACCCATCGGCCACTGGTCCGAGCTGGCTCTCTTGGCTCCCACTGCTCAGAGACTCTGCGAGGG	1933
DB	1753	ACCCATCGGCCACTGGTCCGAGCTGGCTCTCTTGGCTCCCACTGCTCAGAGACTCTGCGAGGG	1813
QY	1933	GGAGTATCTCATAGAGTGGCTCTGCGACCAACCCCAAGCCATCGCATGGAGGGGAGATTCT	1993
DB	1813	GGAGTATCTCATAGAGTGGCTCTGCGACCAACCCCAAGCCATCGCATGGAGGGGAGATTCT	1873
QY	1993	TGTGAGGGCTCCACTCGCATCTCGAGCTCTGCAACAGTCAAGAAATGTCCTCCCGGACAGT	2053
DB	1873	TGTGAGGGCTCCACTCGCATCTCGAGCTCTGCAACAGTCAAGAAATGTCCTCCCGGACAGT	1933
QY	2053	CTTCGACTTCCTGGTCTGCTCAGTGTGCGGAGCACCAACAGACAGCATTTCAAGAGGGCGAC	2113
DB	1933	CTTCGACTTCCTGGTCTGCTCAGTGTGCGGAGCACCAACAGACAGCATTTCAAGAGGGCGAC	1993
QY	2113	TACAGATTCGAGAGCCCTTACACTCTCAGTATGAGAGATACGACATTAATGCAAACTCTACTGTATC	2173
DB	1993	TACAGATTCGAGAGCCCTTACACTCTCAGTATGAGAGATACGACATTAATGCAAACTCTACTGTATC	2053
QY	2173	GCACAGGATTTGATTTCTCTTTTCTTGTGCATAATTAAGTCAAGATGGGACTCCATCC	2233
DB	2053	GCACAGGATTTGATTTCTCTTTTCTTGTGCATAATTAAGTCAAGATGGGACTCCATCC	2113
QY	2233	TGCGAGGATAGCCGCTAATGTTTGTATAGATCGGATATGTGAGAGAGTGTGGATGTGAACAAT	2293
DB	2113	TGCGAGGATAGCCGCTAATGTTTGTATAGATCGGATATGTGAGAGAGTGTGGATGTGAACAAT	2173
QY	2293	GTCTCTTGGATCTGTAGCTGCTGTGAAGAGCTCTGTGGGGGTGTGAACGGGAATAACTCAGCC	2353
DB	2173	GTCTCTTGGATCTGTAGCTGCTGTGAAGAGCTCTGTGGGGGTGTGAACGGGAATAACTCAGCC	2233
QY	2353	TGCAACGATTTCAAGGGGTCTTACACCAAGACCAACACACACCAACAGATTTATTCACATG	2413
DB	2233	TGCAACGATTTCAAGGGGTCTTACACCAAGACCAACACACACCAACAGATTTATTCACATG	2293
QY	2413	GTGACCATTTCTCTTCGAGCGGAGATCTCGCATATGTAAATGAAAGCTCTCTACCTCC	2473
DB	2293	GTGACCATTTCTCTTCGAGCGGAGATCTCGCATATGTAAATGAAAGCTCTCTACCTCC	2353
QY	2473	TACATTTCTGTGGCAATGCTCTCAGAGGATCTACTCGAATGGCCACTGGACCGTGGAC	2533
DB	2353	TACATTTCTGTGGCAATGCTCTCAGAGGATCTACTCGAATGGCCACTGGACCGTGGAC	2413
QY	2533	TGGCCCCGCCGGTACAAATTTTTCGGGCACTACTTTTCGACTACAGAGGGTCTCTAATATGAG	2593

RESULT 3

US-10-217-774-5

CS-10-217-774-3
; Sequence 5, Application US/10217774

Sequence 3, Application OS/1021
: Publication No. US20020193583A1; ADDITION NO: 0320020133363A1
; GENERAL INFORMATION:

APPLICANT: Fridde, Carl Johan

APPLICANT: FRIDGATE, CAZI JOHAN
APPLICANT: HILBUN, ERIP

APPLICANT: HILBURN, ERIN
: TITLE OF INVENTION: NO. 118

1. TITLE OF INVENTION: NO. 33

; TITLE OF INVENTION: Same
 : FILE REFERENCE: IFV-0218-1123

FILE REFERENCE: LEX-0219-USA

;; CURRENT APPLICATION NUMBER: US/10/217,777
: CURRENT FILING DATE: 2002-08-10

; CURRENT FILING DATE: 2002-08-12
 ; PRIOR APPLICATION NUMBER: US/00/030 073

; PRIOR APPLICATION NUMBER: US/09/930,872
; PRIOR FILING DATE: 2001-08-14

PRIOR FILING DATE: 2001-08-14

; PRIOR APPLICATION NUMBER: US 60/225,852

; PRIOR FILI

; NUMBER OF SEQ

; SOFTWARE: Fas

; SEQ ID NO 5

; LENGTH: 404

Query Match

Query MacCn	64.8%	score 18/4.6;	DB 13;	Length 4042;
Best Local Similarity	84.4%	Pred NC 0.		

Best Local Similarity 84.4%; Pred. No. 0;
Matches 2287; Conservative 0; Mismatches

202 GCAGGCACCTGGGTGGGCATGGGACCCGGCAGCGGCCTGGGCGCCGCGCTCC

202 SCAGGCACCTGCTGTGCGCCATGGGACCCGACGGCAGGCTGGGAGCCCGAGCTCCC 281

167 GCAGGCACTGGCTGGGCAATGGACCTGGACGGCCAGCCCTGGCTCC

167 GCAGGCACCTGCGTGGCAATGGACCCGCAGCGGAGCGCCCTGGGAGCCCGAGCGTCCC 226

262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000

262 GCGTCCCTCCACCGCGGAGCGCCGGCTGGATGGAAAAGGCGGGACATGGATGAA 321

[illegible]

22 / GGGTCCCTCCACCCGGGAGCGGCCGGGCTGGATGGAAAGGGCG----- 273

[illegible]

322 GCTGGAAACCATCGTTCTCAGCAAACTAACACAGGAACAGAAACCAACTGCATGTT 381

274 273

202


```
QY 2653 CAGGGAAGAACCGGGTGTGCTTGGGAATACATCTCCATGCTCTGCGGAGCGGAGAG 2712
Db 2631 CAGGGAAGAACCGGGTGTGCTTGGGAATACATCTCCATGCTCTGCGGAGCGGAGAG 2690
QY 2713 CAGCCCTCTGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 2772
Db 2691 CAGCCCTCTGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 2750
QY 2773 TCGGAGCGGGG 2783
Db 2751 TCGGAGCGGGG 2761

RESULT 4
US-10-240-545A-1
; Sequence 1, Application US/10240545A
; Publication No. US20030185828A1
; GENERAL INFORMATION:
; APPLICANT: Tamarouchi Pharmaceutical Co., Ltd.
; FILER: Tamarouchi Pharmaceutical Co., Ltd.
; FILER REFERENCE: 0989-2003
; CURRENT APPLICATION NUMBER: US/10/240-545A
; CURRENT PILING DATE: 2002-10-02
; PRIOR PILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: JP 2000-384300
; PRIOR PILING DATE: 2000-12-18
; NUMBER OF SEQ ID NOS: 26
; SEQ ID NO: 1
; LENGTH: 3666
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3666)
US-10-240-545A-1

Query Match 24.58; Score 708; DB 12; Length 3666;
E-Value Similarity 6.93; Free No. 2;
Matches 1080; Conservative 0; Mismatches 580; Indels 4; Gaps 2;

QY 1124 AGACCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1183
Db 1014 AAAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1073
QY 1184 CTTAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1243
Db 1074 CCATCATGCGAGACGAGTCTCTGAAATAGTTTGGTGGTGGTGGTGGTGGTGGTGGTGG 1132
QY 1244 AGATGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1303
Db 1133 AGATGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1192
QY 1304 AGATGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1363
Db 1193 AGATGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1252
QY 1364 GCAGCTGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1423
Db 1253 GAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1402
QY 1424 CTGGAGACAACTTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1483
Db 1313 CAGGGACAACTTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1372
QY 1484 GCAACATCATGTCCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1543
Db 1373 GCAATATCATGTCCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1432
QY 1544 GCCCGCAGTATCTACAAATTTCTGAGCAGCGCTTGGTGGTGGTGGTGGTGGTGGTGGTGG 1603
Db 1433 GCCCGCAGTATCTACAAATTTCTGAGCAGCGCTTGGTGGTGGTGGTGGTGGTGGTGGTGG 1492
```

```
QY 1604 CAAAGCCTGTGAGGAATACAAAGTATCTCTGAGAAATTTGCCAGGAGAAATTTATATGTCGAA 1663
Db 1493 CCAAGCAAGCAGGAGCAGTATTAATATCTCGGACAACTACAGGACAGATTTATATGTCGTG 1552
QY 1664 ACACAGTGTGAGTGGCAGTTCGGAGAGAAAGCAAGCTCTGCACTGCTGAGACTTTAAAA 1723
Db 1553 ACACAGTGTGAAATGGCAATTTGAGAGAAAGCAAGCTTATGCACTCTGCTGCTGCTGCTG 1612
QY 1724 AGGACATCTGTAAGCCCTGTGTCGCAATCTGATTTGGAAGAAATGTGAGACTTAAATTTA 1783
Db 1613 AGGATATTTGCAATCACTTTGTCGCAAGGAGTGGGCAAGGTGTGAGACCAAGTTA 1672
QY 1784 TGCCAGCAGCAGAGGACAAATTTGTGAGCATGACATGTGTCGCCGGAGGAGCAGTGTG 1843
Db 1673 TGCCCGCAGCAGAGGAGGACCGTTTGTGCTTGTAGTATGTGTCGCGAGGAGCAGTGTG 1732
QY 1844 TGAATATGTGTGATGAAGGCCCOMAGCCCAACCATGCGCACTGTGCGAGCTGTCTTCTT 1903
Db 1733 TNAAGTTTGGGGAGCTCGGGGCCCGGCCCATCCAGCCGAGGTGTGCGGAGTGTGCAAGT 1792
QY 1904 GTTCCCATGTCTCAGGACCTGCGAGGAGGAGTATCTCATAGAGAGTGTGCTCTGCACTG 1963
Db 1793 GTTCAGATGTTCCTCGGACATGTGTGAGAGGAGTCAAGTTCAGAGGAGACACTGTGCAAT 1852
QY 1964 ACCCAGCCCATCGCATGAGAGGAGGAGTCTGTGAGGGTGTGCACTGTGAGCTCTGAGCT 2023
Db 1853 ACCCAGCCCTCAGTATGTGCTTATCTGTCCAGGTTCTAGCCGTATTTATCAGCTGT 1912
QY 2024 GCACAGTCAGAAATGTCCCCGGGACAGTGTGTGACTTCTGCTGTGCTGTGCTGTGCTGTG 2083
Db 1913 GCAATATTAACCTTGCATGAAATAGTCTTGAATTTTCGGGCTCAACAGTGTGCAAGAT 1972
QY 2084 ACACAGCAGCAGATTCAGAGGGGAGCCTACAGTGGAGGCTTTACACTCAAGTAGAGAG 2143
Db 1973 ATACAGCAAACTTTTCGCTGATGTGTTTACAGTGGAGAACCTATACAAAGAGTGGAG 2032
QY 2144 ATCAGACTTATGCAAACTCTACTGTATGTCGAGAGGAGTGTGATTTCTTCTTTCTTGT 2203
Db 2033 AGGAGATGCTGATGCAAACTGTACTGTGCAAGGCTGAGAACCTTTGAAATTTTTTTTGT 2092
QY 2204 CAAATTAAGTCAAAAGTGGAGTCCATGCTCGAGGAGTGGCCGTAAATGTTGTATAGATG 2263
Db 2093 CCGGCAAAAGTGAAGATGGAATCTCTCTCTCCCAACAAATAATGATGTTGTATTTGAGC 2152
QY 2264 GATATGTGAGAGTGTGATGTGCAATGCTCTTGGATCTGATGCTGTGTGAGAGCTCT 2323
Db 2153 GGGTTTGTGAATAGTGTGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 2212
QY 2324 GTGGGTGTGTAAGGGAATTAATCTAGCTGCTGCAATGATGATGATGATGATGATGATGATG 2383
Db 2213 GTGGGTGTGTAAGGGAATTAATCACTTGAAGTGTGATGATGATGATGATGATGATGATGATG 2272
QY 2384 ACACACACACACACACACACACACACACACACACACACACACACACACACACACACACAC 2443
Db 2273 AGCTTAAGCAAAATGAATATTTATGCTGAGTGTGATGATGATGATGATGATGATGATGATG 2332
QY 2444 GCATCTATGAATGAAGCTCTTACCTCTTACATTTCTGAGGAGGAGTGTGATGATGATG 2503
Db 2333 AATTCGAGAGTGTGAGGTTCTTCTGAGTGTGATGATGATGATGATGATGATGATGATGATG 2392
QY 2504 ACTTCTGATGAGGCTGTGAGCCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2563
Db 2393 ATTCTCTCTGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2452
QY 2564 CTTTGTGATGAGGCTGTGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2623
Db 2453 CGTTTGAATACAGGCTCTTTTCAACCCCGGAGAGCTGTGATGAGGAGGAGGAGGAGGAGGAG 2512
QY 2624 ACAGACATGATGATGAGGAGTGTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2683
Db 2513 ATGAGACGCTGGTCTTTTGAATTTCTGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2572
```


US-10-226-560-3
 ; Sequence 3, Application US/10226560
 ; Publication No. US20030050464A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hu, Yi
 ; APPLICANT: Nepomichy, Boris
 ; APPLICANT: Walke, D. Wade
 ; APPLICANT: Friddle, Carl Johan
 ; FILE OF INVENTION: No. US20030050464A1 Human Proteases and Polynucleotides Encoding
 ; TITLE REFERENCE: LEX-0372-USA
 ; CURRENT FILING DATE: 2002-08-22
 ; PRIOR FILING DATE: 2001-08-22
 ; PRIOR FILING DATE: 2001-07-27
 ; PRIOR FILING DATE: 2000-07-28
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 3
 ; LENGTH: 3013
 ; TYPE: DNA
 ; ORGANISM: homo sapiens
 ; US-10-226-560-3

Query Match 22.38; Score 673.4; DB 14; Length 3013;
 Best Local Similarity 65.61; Pred. No. 2.1e-202;
 Matches 997; Conservative 0; Mismatches 521; Indels 1; Gaps 1;

QY	1124	AGACCTGGTGGTGGTCGCAAAAGATGATGCAAAACATGCGCCATGAAATATACCA	1183
DB	1397	AAAGTGGTGGTGGAGCTGATCTCTGGAACAGACCTGGAGGATTAATGATCAA	1456
QY	1184	CTACGCTCTGACATCTCAACATGATCTGCTTTTCAAGATGATGATGGGA	1243
DB	1457	CAATATGACGACGCTCTGATAGTTTGG-TCATGGCAGTCTGCGCCATTTGGAA	1515
QY	1244	AGAGTGGGACTGTCATGACACGCGCTCTACTGACTGCTCGATATATGTTCTCTGA	1303
DB	1516	AGATGGCAGACATGATATGCCATCTACTACACGATTTGATTTGTTCTTGA	1575
QY	1304	AGATAGGCTGATCTGGATTTGCCATCCATGATGGAATGTGTAGTAATATC	1363
DB	1576	AGATAGAACATGTGACCTCTAGGTTTGCCCCATCAGTGGAAATGTGCTTAAGTACC	1635
QY	1364	GGAGCTGACAGATTAATGAGATACAGCTCTTGACCTGCGCTTACCATGCGCATGAT	1423
DB	1636	GAATGTGACATCATAGGACACAGGATGCGCTTCCCTTACCATGCTCATGAT	1695
QY	1424	CTGGACACATCTTGGCAGTATCTGATGAGAGGAGCAATGTGTAAAGATCGCAGG	1483
DB	1696	CAGGACACATCTTGGTATGATCTACGAGAGGAGATCCCTCGAGAAAGCTGAG	1755
QY	1484	GCACATATCTCTCCATCTGATGAGGAGCAATGAGTCTCTCTGGTCACTTGA	1543
DB	1756	GCAATATCATGCTCCCACTGACGGAACATGATGTTTATGCTTCTTCTTGA	1815
QY	1544	CGCCGAGTATCTACAAATTTCTGACGACCGCTCATGCTATCTGCTTCTCATCAGC	1603
DB	1816	GCCTGAGTATCTCAAGAAATCTCTCAGCACCTTCAGCGGGGTGTCTAGTGGTGGAC	1875
QY	1604	CAAGCTCTGAGGATACAGTATCTCGAAGTTCAGAGGAAATATATATGTCGA	1663
DB	1876	CAAGCAAGCAGGACAGTATAATATCTCGACAAATCTACAGGACAGATTATGATGCTG	1935
QY	1664	ACACAGTGGAGTGGATTTGAGAGAGCAAGCTCTGCTGCTGAGCTTTAAA	1723
DB	1936	ACACAGTGTAAATGCAATTTGGAGCAAGCAAGTATGACCTTGTGTTGTA	1995
QY	1724	AGGACATCTGTAAAGCCCTGTGGTCCATCTATTTGAGAGAAATGTGACATTTTA	1783
DB	1996	AGGATATTTGCAATCACTTTGTTGGCCAGGCTAGGCCACAGGTGTGAGACCAAGTTTA	2055

QY	1784	TGCCAGCAGCAAGAGCAAAATTTTGGGCATGACATGTGGTGGCGGAGCAGATGTTG	1843
DB	2056	TGCCCGCAGCAGAAAGGACCGTTTGTGGCTGTAGTATGTGGTGTGCGCAAGCCAGTGC	2115
QY	1844	TGAAATATGGTATGAGAGCCCAAGCCCAACCCCATGGCCACTTGGTGGAGCTGTCTT	1903
DB	2116	TAAAGTTTGGGAGCTCGGGGCCCGCCCATCCAGCGCCAGTGGTCCGCTGGTGAAGT	2175
QY	1904	GSTCCCATGTCTCCAGAGACTCGGAGGGGAGATCTCATAGAGTGCCTCTTGCAACA	1963
DB	2176	GCTCAGATGTTTCCCGGCAATGTGGTGGAGGAGTCAAGTTTCAGGAGACACATGCAAT	2235
QY	1964	ACCCCAAGCCATCGCATGAGGAGGAGTCTGTGAGGGCTCCACTCGCACTCTGAAGCT	2023
DB	2236	ACCCCAAGCCCTCAGTATGCTGCTTATCTGTCCAGGTTCTAGCCGTTTATCAGCTGT	2295
QY	2024	GCAACAGTCAAGATGTCCCCGGGACAGTGTGACTTCCGTGCTGTCAAGTGCCTGAGC	2083
DB	2296	GCAATATTAAACCTTGCATGAAAAATAGCTTGGATTTTCGGGCTCAACAGTGTCAAGT	2355
QY	2084	ACAACAGCAGACGATTCAGAGGGCGGCACCTACAAGTGAAGCTTTACACTCAAGTAGAAG	2143
DB	2356	ATAACAGCAAAACCTTTCCGTGGATGGTCTTACAGTGGAAACCTATACAAAGTGAAG	2415
QY	2144	ATCAGGACTTATGCAAACTCTACTGTATGCGCAGAGGATTTGATTTCTTTCTTTCTGT	2203
DB	2416	AGGAGATCGATGCAAACTGTACTGTGCAAGGCTGAGAACTTTGAAATTTTTTTTCAATGT	2475
QY	2204	CAAAATAAGTCAAGATGGGACTCCATGCTCGAGGATAGCCGTAATGTTTGTATAGATG	2263
DB	2476	CCGSCAAAGTGAAGATGGAACCTCCCTGCTCCCAAAACAAAATGATGTTGTATTGAGC	2535
QY	2264	GGATATGTGAGAGAGTGGATGTGCAACTGTCTTGGATCTGTGATGCTGTGAGAGCTGT	2323
DB	2536	GGSTTTGTGAACAGTGGGATGTGATCACTGAACCTTAGGCTCTAAAGCAGTTTCAGATGCT	2595
QY	2324	GTGGGGTGTGAACGGGAATAACTCAGGCTGCAAGATCAGAGGGTCTCTACACCAAGC	2383
DB	2596	GTGGCTTTTGAAGGTGATAATCACTTGAAGTTTATTAAGGCTCTGATCTCAACC	2655
QY	2384	ACCACCAACCAACAGTATTATCACTGATGCTGATCCATCTCTTGTGAGCGCGGATCC	2443
DB	2656	AGCATAAAGCAAAATGAATATATTCGGTGGTCTCTCATTCAGCTGCGCCCGGAGCATCG	2715
QY	2444	GCATCTATGAAATGAACCTCTCTACCTCTACATTTCTGTGGCAATGCGCTCAGAGGT	2503
DB	2716	AAATCCAGAGCTGCGAGGTTTCTCTCCAGTTACCTCCAGTTTCGAGCTCTCAGTCAAACT	2775
QY	2504	ACTACTGATATGGGCACTGGAACCTGGAAGTGGCCCGCGGTACAAAATTTTGGGCACTA	2563
DB	2776	ATTACTCAACCGGGGCTGGAGCATCGACTGGCTTGGGAGTTTCCCTTCGCTGGGACCA	2835
QY	2564	CTTTGCACTACAGACGCTCTTAATGAGCCCGAGAACTTAATCGCTACTGGACCAACA	2623
DB	2836	CGTTTGAATACAGCGCTCTTTTCAACCGCCCGGAAGCTGTGTACCGCGCAGGCTCCACA	2895
QY	2624	ACGAGACATGATTTGGA	2642
DB	2896	ATGAGACGCTGGTCTTTGA	2914

RESULT 7

US-10-217-774-1

; Sequence 1, Application US/1021774
 ; Publication No. US20020193583A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Friddle, Carl Johan
 ; APPLICANT: Hilburn, Erin
 ; FILE OF INVENTION: No. US20020193583A1 Human Proteases and Polynucleotides Encoding
 ; TITLE REFERENCE: LEX-0219-USA
 ; FILE REFERENCE: LEX-0219-USA
 ; CURRENT APPLICATION NUMBER: US/10/217,774

	Query Match Best Local Matches 1079;	Similarity 71.5%;	Score 663;	DB 13;	Length 1476;	Mismatches 185;	Indels 245;	Gaps 6;
QY	202	GCAGGCACCTCGCTGGCCATGCGACCCGCGACGGCAGCGCTGGGAGCCCGAGCGGTCCC	261					
DB	69	GCAGGCACCTCGCTGGCCATGCGACCCGCGACGGCAGCGCTGGGAGCCCGAGCGGTCCC	128					
QY	262	GGGTCTCTCTCACCCGCGAGCGCGCGGCTGTGATGGAAAGGGGCGGACATGGATGAA	321					
DB	129	GGGTCTCTCTCACCCGCGAGCGCGCGGCTGTGATGGAAAGGGGCGGACATGGATGAA	175					
QY	322	GCTGGAAACCATCTGTTCTCAGCAAACTAACAACGAGACAGAAACCAACCAACTGCATGTT	381					
DB	176	-----	175					
QY	382	CTCACTCAATATGACTTGGTCTCTGCTCCTACGAGGTTGACACAGGGGCGATTACGTGTCC	441					
DB	176	-----AATATGACCTTGGTCTCTGCTCCTACGAGGTTGACACAGGGGCGATTACGTGTCC	228					
QY	442	CATGAATCATGCAACATCAGCGCGCGAGAGAGCACTGCGCGTGTCCGAGGTTGAGTCT	501					
DB	229	CATGAATCATGCAACATCAGCGCGCGAGAGAGCACTGCGCGTGTCCGAGGTTGAGTCT	288					
QY	502	CTTCACCTTCGGCTGAAAGCGCCGACGCAACGATTTCCACATGGATCTGAGCACTTCGACG	561					
DB	289	CTTCACCTTCGGCTGAAAGCGCCGACGCAACGATTTCCACATGGATCTGAGCACTTCGACG	348					
QY	562	AGCCTAGTGGCTCTGGCTTTATGTGTGACAGGTTGGGAAAGACAGGSCATGAGTGTGTG	621					
DB	349	AGCCTAGTGGCTCTGGCTTTATGTGTGACAGGTTGGGAAAGACAGGSCATGAGTGTGTG	408					
QY	622	CAGCTATTACCGCAGAGGACTCTGTGTTTGTATGACGCTTGGGAAAGACAGGACAGAC	681					
DB	409	CNAGCTTTPACCCACAGAGAGCTCTGTGTTTCTATCAAGTCTTTGCGATCACAAGGAC	468					
QY	682	TGCGCATCCGATGGAGGGAAGTTCTGTGAGGCGTCCACTGCGCATCTGAGCTCTGCAAC	741					
DB	469	TCTTCAGTGG-----CCCTTTCAACTGCCCAAGGCTGTGCA	504					
QY	742	AGTCAGAAATGTCCTCCGGGACAGTGTGATCTTCGTCGTCGCTCAGTGTGCGGACGACAC	801					
DB	505	GGCATGATACGAACAAGAGGACAGATTACTTCTAAGGCGACTTCCTTTCACACTCTCA	564					
QY	802	AGCAGACGATTACAGAGGCGCGCACTACACAGTGG--AAGCTTTACACTCAAGTAGAGCGCG	859					
DB	565	TGGAACTCGGCGAGCTGCCAAGGCGAGCTGCCATCCGACTGTACAAGAGATCC	624					
QY	860	ACTTATTCGAAACTCTACTGTATGCGAGAAGGATTGATTCTCTTTCTTTTGTTCGAAATA	919					
DB	625	ACAGAGCCCATGCTCTCTGGGCGCAGTGAAGTCTCTGTGTGACCTC-----AAGG	672					
QY	920	AAGTCAAAAGATGGGACTCCATGCTCGAGGAGTAGCGGTAACTTTGTATAGATGGGATAT	979					
DB	673	ACATGGGAGCTGGACACATCAACCCCTCCACACAGACGACCTTCGCGCTGGGACTGCCACA	732					
QY	980	GTGAGCTCAGTGTGTGTCACATCTCTCGCG--ACATGCGCCCGAGCTTCCCAAGGAGACCTC	1038					
DB	733	AAGCAGCAATTTCTGTGGAGAGCGCAAGAAATACATGCGCCAGCTCTCCAGAGAGACCTC	792					


```

; GENERAL INFORMATION:
; APPLICANT: Bandaru, Rajasekhar
; TITLE OF INVENTION: 65577, A Human Matrix
; FILE OF INVENTION: Metalloproteinase and Uses Therefor
; FILE REFERENCE: MP101-049PIRNM
; CURRENT APPLICATION NUMBER: US/10/103,377C
; CURRENT FILING DATE: 2002-08-26
; PRIOR APPLICATION NUMBER: 60/278,347
; PRIOR FILING DATE: 2001-03-23
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 2052
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2055)
; US-10-103-377C-3

Query Match          9.2%; Score 265.8; DB 14; Length 2052;
Best Local Similarity 62.0%; Pred. No. 5.9e-73;
Matches 420; Conservative 0; Mismatches 257; Indels 0; Gaps 0;

QY 1966 CCAAGCCTGCGATCGAGGAGAGTCTGTGAGGCTCCACTGCGACTCGAGCTGCG 2025
DB 1321 CACAGCCTGAGTATGGTGTATCTGCTCAGGTTTACCGGTATTATCAGCTGTC 1380
QY 2026 CACAGCTGAATGCTCCGGGAGCTGTGACTTCCTGCTGCTCAGTGTGCGAGC 2085
DB 1381 AATATTACCTTGCATGAATAGCTTGAATTTCCGGCTCAACAGTGTGAGNAT 1440
QY 2086 CACAGCAGCATGTACAGGCGGACCTACAGTGTGAGGCTTACAGTGAAGAT 2145
DB 1441 AACAGCAACCTTTCCGTGATGATTCACGCTGAGAACCTTACAAAGTGTGAG 1500
QY 2146 CAGGCTTATGCAACCTACTGTATCGCAGAGAGATTTGATTTCTTTCTTTC 2205
DB 1501 CAGAGCTGTCGACCTGCTGACAGCTGAGACTTGAATTTTCTTTCGAAATG 1560
QY 2206 AATAAGAGTCAAGATCGGACTTCACTGCTCGAGGATAGCGCTAATGCTGTA 2265
DB 1561 GGCAGAGTGAAGATCGAAGTCTCTCTCCCAACAAACAAATATGATTTGTA 1620
QY 2266 ATATGTGAGAGATGTGATGTGCAATGTCTTGGATCTGATCTGTTGAAGCT 2325
DB 1621 GTTGTGAAGTGTGAGTGTGATGATGATGATGATGATGATGATGATGATGAT 1680
QY 2326 GGGGTGTGTAACGGGAATTAATCAGCTGCAACATTCACAGGGGTCTTACAC 2385
DB 1681 GGGGTGTGTAACGGGTATTAATCACTTGAAGTGTGTAACGGGTCTTACAC 1740
QY 2386 CACACACACACAGATTTATCATATGCTGCAATTCCTTCTGGAGCTCGAGT 2445
DB 1741 CATAAAGCAAAATGAATATATATCCGGTGTCTCATTTCCAGCTGCGCGC 1800
QY 2446 ATCTATGAATGAACCTCTCTACCTCTACATTTCTGTGCGCAATGCCCT 2505
DB 1801 ATCCAGAGCTGCGAGGTCTTCTCCAGTATACCTGSCAGTTCGAAGCTCAG 1860
QY 2506 TACCTGAATGGCCTGAGACCGTGTGAGCTGCGCGCGGTACAAATTTTCC 2565
DB 1861 TACCTCAGCGGGGCTGGAGCATGAGCTGGCTGTGGGAGTTCCTCCCTT 1920
QY 2566 TTGACTACAGAGCTCTATATAGAGCCGAGAGACTTAATCTGCTACTGGA 2625
DB 1921 TTGGAATACAGAGCTCTTTTCAACCGCCCGGAAAGCTGTATACGCGC 1980
QY 2626 GAGACACTGATTGTGGA 2642
DB 1981 GAGACACTGATTGTGGA 1997

RESULT 12
US-10-103-377C-1
; Sequence 1, Application US/10103377C
; Publication No. US20030073098A1
; GENERAL INFORMATION:
; APPLICANT: Bandaru, Rajasekhar
; TITLE OF INVENTION: 65577, A Human Matrix
; FILE OF INVENTION: Metalloproteinase and Uses Therefor
; FILE REFERENCE: MP101-049PIRNM
; CURRENT APPLICATION NUMBER: US/10/103,377C
; CURRENT FILING DATE: 2002-08-26
; PRIOR APPLICATION NUMBER: 60/278,347
; PRIOR FILING DATE: 2001-03-23
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2377
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (323)...(2374)
; US-10-103-377C-1

Query Match          9.2%; Score 265.8; DB 14; Length 2377;
Best Local Similarity 62.0%; Pred. No. 8.1e-73;
Matches 420; Conservative 0; Mismatches 257; Indels 0; Gaps 0;

QY 1966 CCGAGCCTGCGATCGAGGAGAGTCTGTGAGGCTCCACTGCGACTCGAGCTG 2025
DB 1643 CACAGCCTCAGTATGGTGTATCTGCTCAGGTTTACCGGTATTATCAGCTGTC 1702
QY 2026 CACAGTCAAGATGCTCCGGGAGCTGTGACTTCCTGCTGCTCAGTGTGCGAGC 2085
DB 1703 AATATTACCTTGCATGAATAGCTTGAATTTCCGGCTCAACAGTGTGAGNAT 1762
QY 2086 CACAGCAGCATGTACAGGCGGACCTACAGTGTGAGGCTTACCTGAAGAT 2145
DB 1763 CACAGCAACCTTTCCGTGATGATTCACAGTGTGAGAACCTTACAAAGTGTG 1822
QY 2146 CAGGCTTATGCAACCTACTGTATCGCAGAGAGATTTGATTTCTTTCTTTC 2205
DB 1823 GAAGTCGAGTCAAGCTGCTGCTGCAAGCTGAGAACTTTTGAATTTTTCGAA 1882
QY 2206 AATAAGTCAAGATCGGACTTCACTGCTCGAGGATAGCGCTAATGTTTGTAT 2265
DB 1883 GGCAGAGTGAAGATCGAAGTCTCTCTCCCAACAAACAAATATGATTTTGT 1942
QY 2266 ATATGTGAGAGATGTGATGTGCAATGTCTTGGATCTGATCTGTGAGAGCTCT 2325
DB 1943 GTTGTGAAGTGTGAGTGTGATGATGATGATGATGATGATGATGATGATGAT 2002
QY 2326 GGGGTGTGTAACGGGAATTAATCAGCTGCAACATTCAGCTGAGGGTCTTAC 2385
DB 2003 GGGGTGTGTAACGGGTATTAATCACTTGAAGTGTGTAACGGGTCTTAC 2062
QY 2386 CACACACACACAGATTTATCATATGCTGCAACATTCCTTCTGGAGCTCG 2445
DB 2063 CATAAAGCAAAATGAATATATATCCGGTGTCTCATTTCCAGCTGCGCGC 2122
QY 2446 ATCTATGAATGAACCTCTCTACCTCTACATTTCTGTGCGCAATGCCCT 2505
DB 2123 ATCCAGAGCTGCGAGGTCTTCTCCAGTATACCTGCGAGTTCGAAGCTCAG 2182
QY 2506 TACCTGAATGGCCTGAGACCGTGTGAGCTGCGCGCGGTACAAATTTTCC 2565
DB 2183 TACCTCAGCGGGGCTGGAGCATGAGCTGGCTGTGGGAGTTCCTCCCTT 2242
QY 2566 TTGACTACAGAGCTCTATATAGAGCCGAGAGACTTAATCTGCTACTGGA 2625
DB 2243 TTGGAATACAGAGCTCTTTTCAACCGCCCGGAAAGCTGTATACGCGC 2302
QY 2626 GAGACACTGATTGTGGA 2642
```


Wed Oct 29 15:41:41 2003

1871	Db		GTGCGAGCTTCGCATATGTCCTTTCCGAGGAAGATTTATTAACTCGAAACCCCTATACT-	1929
2135	Qy		AGGTGAGATCAGGACTTATGCAATCTCTCACTGTATCGAGAGGATTTGATTTCTCT	2194
1930	Db		--GGGTGGGGTAAACCTTGTGCATTAACTGCTTGCCTGAAGCTATTAATTTCTACA	1987
2195	Qy		TTCTTTGTCAATTAAGTCAAAAGTGGCACTCCCATGCTCGSAGATAGCGTAATGTTT	2254
1988	Db		CTGAACTGCTCCTCGGCTGATCGATGGGACCCGAGTGCNAATCGGATTCACCTGCAATCT	2047
2255	Qy		GTATAGATCGGATATGTAGAGAGTGTGATGTGCAATGTCCTTTGGATCTGATGCTGTG	2314
2048	Db		GCATCAATGGAGNATGCAAGCNCGTAGGCTGTGATTAATTTTGGATCTGATGCTAGGG	2107
2315	Qy		PAGAGCTCTGTGGGGTGTCTAACCGGGAATAACTCAGCGCTGCACGATTCACAGGGGCTCT	2374
2108	Db		AGATAGATGTGCAATCTGTGGAGGGGACGGAAACATGTGATGTCATTTGAAAGGTTCT	2167
2375	Qy		ACNCCAGCACCACCAACACAGTATTATACATGTGTCAACATTCTCTCTGAGGCC	2434
2168	Db		TCAATGATTCACTGCCCGCGGGGGCTTACATGGAAGTGTGTGCAGATACCAAGAGGCTCTG	2227
2435	Qy		GGAGTATCCGCACTATATGAATAGAACGTCTTACCTCTTACATTTCTGTGCGCAATGCC	2494
2228	Db		TTCACTTGAAGTTAGAGAAGTTGTCATGTCAAGAAGAACTATATTGCTTTTAAATCTGAAG	2287
2495	Qy		TCAGAAGGTACTTACTCGATGGGCATCTGGACCGTGGACTGCGCCGCGGTACAAATTTT	2554
2288	Db		GAGATGATTACTATATTATATGCTGCTCGGACTATTGACTGCGCTAGGAATTTTGATGTTG	2347
2555	Qy		CGGCACTACTTTTGGACTACAGACGCTCTATATTAATGAGCCGAGAACTTAATCTGCTACTG	2614
2348	Db		CTGGGACAGCTTTTCAATTACAGAGACCAACTGTATGAACGAATTCCTTTGGAAGCTTAG	2407
2615	Qy		GACCAACCAACAGACACTGATTGTGGAGCTCTGTTTTCAGGGAAGGAACCGCGGTGTT	2673
2408	Db		GTCTTACTTCAGAAATCTCATGCTCATGGTCTGCTTCAGAACAGATATTTGGGAATT	2466

Search completed: October 28, 2003, 23:39:24
Job time : 764 secs

ALIGNMENTS

ABG76897
ID ABG76897 standard: Protein: 952 AA.

05-NOV-2002 (first entry)

DE Human ADAM-TS 7-like protein #2.

Human; NOVX; cardiomyopathy; atherosclerosis; cell signal processing;
breast cancer; Alzheimer's disease; epilepsy; Huntington's disease;
anxiety; behavioural disorder; multiple sclerosis; myasthenia gravis;
neurodegeneration; Parkinson's disease; pain; stroke; endometriosis;
autoimmune disease; allergy; addiction; asthma; transplantation;
graft versus host disease; systemic lupus erythematosus; scleroderma;
psoriasis; Crohn's disease; HIV infection; human immunodeficiency virus;
atherosclerosis; cirrhosis; rheumatoid arthritis; diabetes; pancreatitis;
chromocytoma; B-cell lymphoma disorder; metabolic disorder; obesity;
polycystic kidney disease; renal tubular acidosis; skin disorder;
congenital diarrhoea; respiratory disease; gastro-intestinal disease;
muscle disorder; bone disorder; joint disorder; skeletal disorder;
haematopoietic disorder; urinary system disorder; osteoporosis;
dental disease; dental infection; growth disorder; reproductive disorder;
hypogonadism; fertility disorder; viral infection; bacterial infection;
parasitic infection; metabolic pathway modulation; gene therapy;
zinc metalloproteinase; ADAM-7; alpha-2-macroglobulin precursor;
lead sodium/bile acid cotransporter; proinhibin; Wnt/Ca²⁺ signalling;
gag junction beta-5 protein; heparin/plasma transmembrane serine protease.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result	No.	Score	Query		Length	DB	ID	Description
			Match	%				
1	5236	100.0	952	23	ASG76897		Human ADAM-15 7-li	
2	5110	97.6	986	23	ASG76894		Human zinc metallo	
3	4103	78.4	791	23	ASG76895		Human zinc metallo	
4	3585.5	73.7	1324	23	AAU79497		Human ADAMTS prote	
5	3558.5	73.7	1324	23	AAU79497		Human metalloprote	
6	3446.5	73.5	1324	23	AAO15284		Human ADAMTS prote	
7	3446.5	73.5	1324	23	AAO15284		Human ADAMTS prote	
8	3440.5	73.3	1324	24	ASG72431		Novel Human ADAMTS prote	
9	3440.5	73.3	1324	23	AAU80153		Human ADAMTS prote	
10	3750.5	71.6	1170	23	AAU79500		Human ADAMTS prote	

DB 481 LGFAPISGMSKYRSTINEDTGLGTAFTIAHSGHFGWTHOGEWNCKSEGNINSPT 540
 QY 499 LAGRNQVFSNCSQRLHFLSTAQAICLAQPKPVKEYKPEKLFGBELDANTCKWQ 558
 DB 541 LAGRNQVFSNCSQRLHFLSTAQAICLAQPKPVKEYKPEKLFGBELDANTCKWQ 600
 QY 559 FGKAKLCLMDFDKDICKALWCHRIKGRCKTFKMPAAEGTICGHDMWCRGGQCVKYGDG 618
 DB 601 FGKAKLCLMDFDKDICKALWCHRIKGRCKTFKMPAAEGTICGHDMWCRGGQCVKYGDG 660
 QY 619 PXPETHGWSWMSWSPSCSRCTCGGVSRSRLCTNPKSHGKFCGSGTRTLKLNCKP 678
 DB 661 PXPETHGWSWMSWSPSCSRCTCGGVSRSRLCTNPKSHGKFCGSGTRTLKLNCKP 720
 QY 679 RSVDVFAAQCQAEHNSRRFRGRHYKFPYQVVEDOOLCKLYCIAEGDFEFSLNRYKDG 738
 DB 721 RSVDVFAAQCQAEHNSRRFRGRHYKFPYQVVEDOOLCKLYCIAEGDFEFSLNRYKDG 774
 QY 739 TPCSDSRNVYCIDGICRVCQCNVLGSDAVEDVCGVCGNNSACTIHRGLYTKHHHTNY 798
 DB 775 TPCSDSRNVYCIDGICE-XCQCNVLGSDAVEDVCGVCGNNSACTIHRGLYTKHHHTN-H 832
 QY 799 YHNVTPSGARSIRIYENNVSTYSISVNALARYLNGHTVDMQRYFSGTTFDYRS 858
 DB 833 YHNVTPSGARSIRIYENNVSTYSISVNALARYLNGHTVDMQRYFSGTTFDYRS 892
 QY 859 YNEPENLIATGPNETLIVELLFOQRNPGVAMEYSNPRLOTQKPPAQPSYTWAIVRSEC 918
 DB 893 YNEPENLIATGPNETLIVELLFOQRNPGVAMEYSNPRLOTQKPPAQPSYTWAIVRSEC 952
 QY 919 SVSCGGRCLPVLLEAACOPSATYIALAFLES 952
 DB 953 SVSCGGRCLPVLLEAACOPSATYIALAFLES 986

RESULT 3

ABG76895
 ID ABG76895 standard; Protein; 791 AA.
 AC ABG76895;
 DT
 DX
 DZ
 KW Human zinc metalloprotease-like protein #2.
 KW Human; NOVX; cardiomyopathy; atherosclerosis; cell signal processing;
 KW breast cancer; Alzheimer's disease; epilepsy; Huntington's disease;
 KW anxiety; behavioural disorder; multiple sclerosis; myasthenia gravis;
 KW neurodegeneration; Parkinson's disease; pain; stroke; endometriosis;
 KW autoimmune disease; allergy; addiction; asthma; transplantation;
 KW graft versus host disease; systemic lupus erythematosus; scleroderma;
 KW psoriasis; Crohn's disease; HIV infection; human immunodeficiency virus;
 KW atherosclerosis; cirrhosis; rheumatoid arthritis; diabetes; pancreatitis;
 KW thrombocytopenia; bleeding disorder; metabolic disorder; obesity;
 KW glucose transport defect; glomerulonephritis; hypercalcaemia;
 KW polycystic kidney disease; renal tubular acidosis; skin disorder;
 KW congenital diarrhoea; respiratory disease; gastro-intestinal disease;
 KW muscle disorder; bone disorder; joint disorder; skeletal disorder;
 KW haematopoietic disorder; urinary system disorder; osteoporosis;
 KW dental disease; dental infection; growth disorder; reproductive disorder;
 KW hypogonadism; fertility disorder; viral infection; bacterial infection;
 KW parasitic infection; metabolic pathway modulation; gene therapy;
 KW zinc metalloprotease; ADAM-7; alpha-2-macroglobulin precursor;
 KW zinc metalloprotease; ADAM-7; alpha-2-macroglobulin precursor;
 KW macrophage stimulating protein precursor; fatty acid-binding protein;
 KW gap junction beta-5 protein; hepsin/plasma transmembrane serine protease.
 OS Homo sapiens.
 XX
 XX
 PN WQ200233087-A2.
 XX
 PD 25-APR-2002.

XX 17-OCT-2001; 2001WO-US32496.
 XX 17-OCT-2000; 2000US-241040P.
 PR 17-OCT-2000; 2000US-241058P.
 PR 17-OCT-2000; 2000US-241063P.
 PR 17-OCT-2000; 2000US-241243P.
 PR 20-OCT-2000; 2000US-242152P.
 PR 23-OCT-2000; 2000US-242482P.
 PR 23-OCT-2000; 2000US-242611P.
 PR 23-OCT-2000; 2000US-242612P.
 PR 24-OCT-2000; 2000US-242880P.
 PR 24-OCT-2000; 2000US-242881P.
 PR 29-DEC-2000; 2000US-259028P.
 PR 20-FEB-2001; 2001US-269813P.
 PR 25-APR-2001; 2001US-286324P.
 PR 29-MAY-2001; 2001US-294108P.
 PR 09-JUL-2001; 2001US-303698P.
 PR 16-OCT-2001; 2001US-0981151.

(CURA-) CURAGEN CORP.

Edinger S, Gerlach V, MacDougall JR, Malyankar UM, Smithson G;
 Millet I, Peyman JA, Stone DJ, Gunther E, Ellerman K, Shinkets RA;
 Padigaru M, Guo X, Patturajan M, Taupier RJ, Burgess CE;
 Zehrhusen BD, Kekuda R, Spytek KA, Gangolli EA, Fernandes ER;
 Gorman L;

WPI; 2002-590434/63.
 N-FSDB; ABS59324.

Cytoplasmic, nuclear, membrane bound and secreted polypeptides and
 nucleic acids encoding the polypeptides for diagnosing and treating
 e.g. cancer, Alzheimer's disease, cardiomyopathy, metabolic disease and
 diabetes

Claim 1; Page 15; 305pp; English.

The present invention relates to new NOVX (NOVI-10) polypeptides. The
 molecules of the invention are useful for treating or preventing a
 NOVX-associated disorder, such as cardiomyopathy, atherosclerosis, or
 a disorder related to cell signal processing and metabolic pathway
 modulation in humans. NOVX polypeptides, nucleic acids and antibodies
 are useful for treating or preventing disorders or syndromes, including
 breast cancer, Alzheimer's disease, epilepsy, Huntington's disease,
 anxiety, behavioural disorder, multiple sclerosis, myasthenia gravis,
 neurodegeneration, Parkinson's disease, pain, stroke, endometriosis,
 autoimmune disease, allergy, addiction, asthma, transplantation,
 graft versus host disease, systemic lupus erythematosus, scleroderma;
 psoriasis, Crohn's disease, HIV (human immunodeficiency virus) infection,
 atherosclerosis, cirrhosis, rheumatoid arthritis, diabetes, obesity,
 thrombocytopenia, bleeding disorders, metabolic disorders, polycystic
 glucose transport defect, glomerulonephritis, hypercalcaemia, polycystic
 kidney disease, pancreatitis, renal tubular acidosis, skin disorders,
 congenital diarrhoea, respiratory disease, gastro-intestinal diseases,
 muscle, bone, joint and skeletal disorders, haematopoietic disorders,
 urinary system disorders, osteoporosis, dental disease and infection,
 pathologies and reproductive disorders, hypogonadism, fertility, and/or other
 pathologies and disorders, viral, bacterial, or parasitic infections.
 The present amino acid sequence represents a NOVX protein of the

Sequence 791 AA;

Query Match 78.4%; Score 4103.5; DB 23; Length 791;

Best Local Similarity 81.2%; Pred. No. 0;

Matches 774; Conservative 9; Mismatches 7; Indels 163; Gaps 8;

QY 1 MKPRARGWRGLAALWLLAQVAQVSPGRSHQGRNKGSGOLEASPRLLSRGPRRLTAMS 60

DB 1 MKPRARGWRGLAALWLLAQVAQVSPGRSHQGRNKGSGOLEASPRLLSRGPRRLTAMS 60

QY 61 PLFSAGTCVRHGTSGSANEPEPASPSSSTRGAAGLDGKGRDMDEAGNHRSQQTNTGTENQ 120

61 PLFSAAGTCVRRHTASGSAPEPSSSTRGAGLGDGRDDEAGNHASQQTGTGTENQ 120
 121 TLHLVLT-QYDLVSAYVDHRRGYVSHIMHQRRRRRAVAVSESLRLKGPDRPHMD 179
 122 TLHLVLTREYDLVSAYVDHRRGYVSHIMHQRRRRRAVAVSESLRLKGPDRPHMD 180
 180 LRTSSSLVAFGIIVTLTGKTGTSVOTLPEDFCFYQGLSRSHRNSPSHGKFCGSGT 239
 181 LRTSSSLVAFGIIVTLTGKTGTSVOTLPEDFCFYQGLSRSHRNSPSHGKFCGSGT 240
 240 LKLNCSQKPRSDVDFRAQCAHNSRRFRGRHYKWPYQVSEADLCKLYCIAGDFDFF 299
 241 LKLNCSQKPRSDVDFRAQCAHNSRRFRGRHYKWPYQVSEADLCKLYCIAGDFDFF 300
 300 SLGNKVKDGFPCSDSRNVCIDIGICELSVVSTSAHMPKBDLFLPDEYKSLRHKHS 359
 301 SLGNKVKDGFPCSDSRNVCIDIGICELSVVSTSAHMPKBDLFLPDEYKSLRHKHS 351
 360 LRSRHEELNVELTVVVDKQMQHGHENITVTLNMVSALEFQGLMKGKGTROH 419
 352 LRSRHEELNVELTVVVDKQMQHGHENITVTLNMVSALEFQGLMKGKGTROH 403
 420 ALLTGLDSCWKEPDLGTFAPISGKCKYRSTINEDTGLATINSHSNFGL 479
 404 ----- 403
 480 HDGNGWCKKSEGNMPTLAGNGVSVSPCSROYLHKFLSTAQICLADQPKVKEIK 539
 404 -----GNI-NIAIVG----- 412
 540 YPEKLPLGELVDANTCKWQGEKAKLCHLDFKDIKALMCHRGCEKTPAARCTI 599
 413 -----LLEDEQDLCKALMCHRGCEKTPAARCTI 447
 600 CGHDWKCRCGQCVKYVDEGEPKTHGHSDWSSWSPCSRTCCGGVSHRSLCTNPKBSHG 659
 448 CGHDWKCRCGQCVKYVDEGEPKTHGHSDWSSWSPCSRTCCGGVSHRSLCTNPKBSHG 507
 660 KCEGSTRTLKLCNSOKPRSDVDFRAQCAHNSRRFRGRHYKWPYQVSEADLCKLY 719
 508 KCEGSTRTLKLCNSOKPRSDVDFRAQCAHNSRRFRGRHYKWPYQVSEADLCKLY 561
 720 CIAEGFDFPFLSLNKKVKGDTFPCSDSRNVCIDIGICERVGCDNLGSDAEDVCGVCGN 779
 562 CIAEGFDFPFLSLNKKVKGDTFPCSDSRNVCIDIGICERVGCDNLGSDAEDVCGVCGN 619
 780 SACTIHRGLYTHKHTNQYHVMVTPSGARSIRIYEMNVSTYSIVRNALRYLNGHMT 839
 620 SACTIHRGLYTHKHTNQYHVMVTPSGARSIRIYEMNVSTYSIVRNALRYLNGHMT 678
 840 VDMQGRYKFGSTTFFRYSNFPENLITGNETLVELLFQGNFGVAMYSNPLGT 899
 679 VDMQGRYKFGSTTFFRYSNFPENLITGNETLVELLFQGNFGVAMYSNPLGT 738
 900 EKOPPAQPSYTWAIVRSECSVCGGRCPLVLLLEAACQPSATYIALAPLES 952
 739 EKOPPAQPSYTWAIVRSECSVCGGRCPLVLLLEAACQPSATYIALAPLES 791

RESULT 4

AAU79497
 ID AAU79497 standard; Protein; 1224 AA.
 AC AAU79497;
 AC AAU79497;
 DT 15-JUL-2002 (first entry)
 DE Human ADAMTS protein #1.
 DE Human; ADAMTS; cytotatic; antidiabetic; antithrombotic;
 KW antithrombotic; antidiabetic; antithrombotic; antithrombotic;
 KW antithrombotic; antidiabetic; antithrombotic; antithrombotic;
 KW antithrombotic; antidiabetic; antithrombotic; antithrombotic;

KW refractory skin ulcer; gastric ulcer; post-operative healing failure;
 KW repressin-type 2N-metalloproteinase domain; disintegrin-like domain; TSP1;
 KW thrombospondin type 1 domain; sexual cycle; tumour; 5p-syndrome deletion;
 OS chromosome 5p15.2-15.3; Cri-du-chat syndrome.
 OS Homo sapiens.
 XX MO200231163-A1.
 XX 18-APR-2002.
 XX 11-OCT-2001; 2001MO-JP08913.
 XX 11-OCT-2000; 2000JP-0311309.
 XX 02-APR-2001; 2001JP-0102905.
 XX (KAZU-) KAZUSA DNA RES INST FOUND.
 PA (MITS-) MITSUBISHI PHARMA CORP.
 FI Ohara O, Nagase T, Nomura N, Yano K, Murakami K, Yasuda S;
 FI Kanzaki K;
 FI WFI; 2002-372277/40.
 FI N-PSDB; AK94822.
 FI Human brain-originated ADAMTS family polypeptide and encoded gene,
 FI applicable in diagnosis and screening compounds for drug compositions
 FI in treating diseases due to e.g. neovascularisation
 FI Claim 1; Page 145-150; 172pp; Japanese.
 XX The invention relates to a polypeptide belonging to the ADAMTS family is
 CC selected from sequences appearing in AU79497, AU79497 and AU79499
 CC and having with the amino acid sequence of a protein having not less than
 CC 50% homology with the amino acid sequence of the polypeptides or a
 CC polypeptide modified from any of the polypeptides but with some amino
 CC acids deleted, substituted, added or inserted. Also included are the
 CC polynucleotides encoding the polypeptides (or their complementary
 CC strands or variants), a recombinant vector containing any of the
 CC polynucleotides, a transformant which is transformed with the recombinant
 CC vector, producing the polypeptide, protein or peptide by culturing the
 CC transformant, an antibody that can recognize the polypeptide, protein or
 CC peptide and screening compounds to promote or inhibit activity of the
 CC polypeptide or protein, or to promote or inhibit expression of the
 CC polynucleotide, vector, transformant or/and antibody, particularly in
 CC the presence of a test compound for contact before evaluating the
 CC activity by measuring signal changes. The polypeptide and encoded gene
 CC are applicable in diagnosis and screening compounds for drug compositions
 CC in treating diseases due to neovascularisation, diabetic omentopathy,
 CC chronic rheumatoid arthritis, angina, refractory skin and gastric ulcers
 CC and post-operative healing failure, including gene therapy.
 CC The gene encoding such polypeptide has conserved repressin-type 2N-
 CC metalloproteinase domain, disintegrin-like domain and TSP1 (thrombospondin
 CC type 1) domain. Its encoded protein is characterised by high expression
 CC in ovaries, changes in expression dose depending on the sexual cycle, a
 CC decrease in tumour cell and location of the gene on the 5p-syndrome
 CC deletion site on chromosome 5p15.2-15.3 (associated with
 CC Cri-du-chat syndrome). The present sequence represents a ADAMTS
 CC protein of the invention.

XX Sequence 1224 AA;

Query Match

73.7%; Score 3858.5; DB 23; Length 1224;

Best Local Similarity 75.6%; Pred. No. 3.1e-309;

Matches 740; Conservative 24; Mismatches 72; Indels 143; Gaps 13;

QY 1 MKPRAGWRGLAALMLLAQVAQVSPQRSHQNRSGQLEASPRLRGRPRILTAMS 60

DB 1 MKPRAGWRGLAALMLLAQVAQVSPQRSHQNRSGQLEASPRLRGRPRILTAMS 31

QY 61 PLFSAAGTCVRRHTASGSAPEPSSSTRGAGLGDGRDDEAGNHASQQTGTGTENQ 120

DB 32 PAAAA-----PGSPVPRPPPAERPG-----WMEKG-----58
 QY 121 TLHLVLTQYDLVSAVEVDHGRGVYSHHMHQRRRAVAVSEVSHLRLKGRPHDFHMDL 180
 DB 59 -----EYDLVSAVEVDHGRGVYSHHMHQRRRAVAVSEVSHLRLKGRPHDFHMDL 112
 QY 181 RTSSSLVAPGFIVQTLGKTGKSVQTLPPEDFCFYQGSLSHSHNSPSHGKFCESGRTLL 240
 DB 113 RTSSSLVAPGFIVQTLGKTGKSVQTLPPEDFCFYQGSLSHSHNS-SVALSTCQGLSGMI 171
 QY 241 KLCNSQKPRDSVDFRAAQCAEHNSRFRGR-----HYWKPKYTOVZA-DLCKLYCI 291
 DB 172 R-----TEADYFLRPLFSLWKL-GRAAQSSPSHVLYKRSTEPHAPGASEVLVT 222
 QY 292 AEGFOFFS--LSNKKVKGDTFCSDSRNVICIDICELSVVSTSAHMPQPPKEDLFLPDE 349
 DB 223 SRTWELAHQHLSSDLRLGLP---QKQHC-----GRKKYMPQPPKEDLFLPDE 270
 QY 350 YKSLRHKRSLSRSHNEELNVETLVVDKMMQNHGHEMITTYVLTLLNWSALFKD-- 407
 DB 271 YKSLRHKRSLSRSHNEELNVETLVVDKMMQNHGHEMITTYVLTLLNWSALFKDGT 330
 QY 408 -----GLMKDGTGRHDAILLTGLD 427
 DB 331 IGGNINIAVGLTLEDEQGLVISHADHTLSSFCQWQGLMGXGDTGRHDAILLTGLD 390
 QY 428 ICSWNEPCDTLGFAPISGKYSRSCYINEDTGLCLAFTHAHESGHNFGMHGEGNMC 487
 DB 391 ICSWNEPCDTLGFAPISGKYSRSCYINEDTGLCLAFTHAHESGHNFGMHGEGNMC 450
 QY 488 K3SEGNIMSPTLAGRNWFGVSPSCROYLHKFLSTAQAICLADQPKPKYKPKLPGF 547
 DB 451 K3SEGNIMSPTLAGRNWFGVSPSCROYLHKFLSTAQAICLADQPKPKYKPKLPGF 510
 QY 548 LYDANTCKMQRGKAKLMLDQKDKDCKALWCHRIKGCETKFMPPAAEGTICGHDMWR 607
 DB 511 LYDANTCKMQRGKAKLMLDQKDKDCKALWCHRIKGCETKFMPPAAEGTICGHDMWR 570
 QY 608 GGGCVKYDGGPPTPHGWSNWSNPSPCSTRCGGVSHRSLCTNPKPSHGKFCESGSTR 667
 DB 571 GGGCVKYDGGPPTPHGWSNWSNPSPCSTRCGGVSHRSLCTNPKPSHGKFCESGSTR 630
 QY 668 TLKLNQSKQPRDSVDFRAAQCAEHNSRFRGRHYWKPKYTOVSDQDLCKLYCIABGDPF 727
 DB 631 TLKLNQSKQPRDSVDFRAAQCAEHNSRFRGRHYWKPKYTOVSDQDLCKLYCIABGDPF 690
 QY 728 FFSLSNKKVKGDTFCSDSRNVICIDICERVOCDNLVSDAVENKGCNNSACTHRG 787
 DB 691 FFSLSNKKVKGDTFCSDSRNVICIDICERVOCDNLVSDAVENKGCNNSACTHRG 750
 QY 788 LYTKGHHHTNYHMTVTPSGARSIYEMNVSTVSVRNALRYVYLNHGTWDPGRYK 847
 DB 751 LYTKGHHHTNYHMTVTPSGARSIYEMNVSTVSVRNALRYVYLNHGTWDPGRYK 810
 QY 848 FSGTTFDYRSYNEPENLIATGPTNETLIVELLFQGRNPGVAMEYSPRLGTEKOPPAQ 907
 DB 811 FSGTTFDYRSYNEPENLIATGPTNETLIVELLFQGRNPGVAMEYSPRLGTEKOPPAQ 870
 QY 908 SYTMAIVSECSVCGGQ 926
 DB 871 SYTMAIVSECSVCGGQ 889

RESULT 5

AAU72891

ID AAU72891 standard; protein; 1224 AA.

XX AC AAU72891;

XX DT 26-FEB-2002 (first entry)

XX DE Human metalloprotease partial protein sequence #3.

Human; protease; PCR primer; cytosolic; immunomodulator; cardiac; vasotropic; anti-migraine; analgesic; endocrine; neurotropic; tranquilizer; hypertensive; hypotensive; neuroleptic; neuroprotective; anabolic; atorectic; anti-inflammatory; aspartyl protease; cysteine protease; metalloprotease; serine protease; cancer; haematopoietic; breast; lung; prostate; cervical; brain; ovarian; bladder; kidney; pain; immune-related disease; cardiovascular disease; neuronal disease; migraine; sexual dysfunction; mood disorder; attention disorder; cognition disorder; hypotension; hypertension; psychotropic disorder; dyskinesia; metabolic disorder; inflammatory disorder.
 OS Homo sapiens.
 XX W0200183782-A2.
 XX 08-NOV-2001.
 XX 04-MAY-2001; 2001WO-US14431.
 XX 04-MAY-2000; 2000US-201879P.
 XX (SUBG-) SUGEN INC.
 XX Plowman GD, Whyte D, Sudarsanam S, Manning G, Caenepeel S; Payne V; WPI; 2002-041502/05.
 XX N-PSDB; AAS97174.
 XX Novel protease polypeptide useful for screening for substances that may be used to treat, e.g., cancers, immune-related diseases, cardiovascular disease, migraine, pain, psychotic and inflammatory disorders -
 XX Claim 28; Figure 2D; 232pp; English.
 XX The invention relates to an isolated, enriched, or purified protease polypeptide (I) and polynucleotide (II) encoding (I). (I) may be used to screen for substances (S) that may modulate its activity. Administering S (which modulates protease activity in vitro) may be used to treat a disease or disorder selected from cancers (e.g., of tissues, of blood or haematopoietic origin, of the breast, colon, lung, prostate, cervical, brain, ovarian, bladder or kidney), immune-related diseases and disorders, cardiovascular disease, brain or neuronal-associated diseases (e.g., central or peripheral nervous system diseases, migraine, pain, sexual dysfunction, mood disorders, attention disorders, cognition disorders, hypotension, hypertension, psychotic disorders, neurological disorders and dyskinesias), metabolic disorders and inflammatory disorders. (I) may also be useful as a diagnostic tool for a disease or disorder such as those above. AAU72891-AAU72910 represent human protease amino acid sequences of the invention.
 SQ Sequence 1224 AA;
 Query Match 73.7%; Score 3858.5; DB 23; Length 1224;
 Best Local Similarity 75.6%; Prod No. 3:4e-309;
 Matches 740; Conservative 24; Mismatches 72; Indels 143; Gaps 13;
 QY 1 MKPRAEGRGLAALWLLAQVAQVSPGSHQGRNGSGQLASPRLLSGRPLRTAMS 60
 DB 1 MKPRAEGRGLAALWLLAQVAQVSPGSHQGRNGSGQLASPRLLSGRPLRTAMS 60
 QY 61 PLESAGTCVRHGTRSGSAWEPEPASPSTRGAGLDGKGRMDRAGNHRSGQTNTGTENQ 120
 DB 32 PAAAA-----PGSPVPRPPPAERPG-----WMEKG-----58
 QY 121 TLHLVLTQYDLVSAVEVDHGRGVYSHHMHQRRRAVAVSEVSHLRLKGRPHDFHMDL 180
 DB 59 -----EYDLVSAVEVDHGRGVYSHHMHQRRRAVAVSEVSHLRLKGRPHDFHMDL 112
 QY 181 RTSSSLVAPGFIVQTLGKTGKSVQTLPPEDFCFYQGSLSHSHNSPSHGKFCESGRTLL 240
 DB 113 RTSSSLVAPGFIVQTLGKTGKSVQTLPPEDFCFYQGSLSHSHNS-SVALSTCQGLSGMI 171

QY 241 KLCNSQKPRDSVDFRAACACAHNSRFRGR-----HYKKPYTOVEA-DLCKLYCI 291
 DB 172 R-----TEADYFLRLPPLSHLWKL-GRAAQSSPSHVLYKRSTEPHAPGASEVLVT 222
 QY 292 AEGDFPFS--LSNKVKGDTPTCSERDNRVCIIGICELSVSTSAHMPQPKEDFLPDE 349
 DB 223 SRTWELAHQPLHSDLRGLP---QKQHF-----GRRKYMPPQPKEDFLPDE 270
 QY 350 YKSLRHRKSLRSHNEELNVTIYVVDKMMQNHENITTYVLTILNWSALFKD-- 407
 DB 271 YKSLRHRKSLRSHNEELNVTIYVVDKMMQNHENITTYVLTILNWSALFKDGT 330
 QY 408 -----GLMGKDGTRHDHAILLTGLD 427
 DB 331 ICGNINIAVLGLLEDEQPLVISHADHTLSSFCQMSGLMGKDGTRHDHAILLTGLD 390
 QY 428 ICSWKNEPCDTLGFAPISGMSKYRSTINEDTGLGLAFTTAHESGHNFQMHDEGNMC 487
 DB 391 ICSWKNEPCDTLGFAPISGMSKYRSTINEDTGLGLAFTTAHESGHNFQMHDEGNMC 450
 QY 488 KKSEGNIMSPTLAGRNGVFSWSPCSRQYLHKFLSTAQAICLADQPKPKYKPKLPG 547
 DB 451 KKSEGNIMSPTLAGRNGVFSWSPCSRQYLHKFLSTAQAICLADQPKPKYKPKLPG 510
 QY 548 LYDANTQCKWQFGEKAKLMLDKFKDICKALWCHIRGRKCKTFMPAAEGTICGHDWMC 607
 DB 511 LYDANTQCKWQFGEKAKLMLDKFKDICKALWCHIRGRKCKTFMPAAEGTICGHDWMC 570
 QY 608 GGQCVKYDGPKPTHGHWSWSSWSPCSRQYVSHRSLCTNPKPSHGKFCGSTR 667
 DB 571 GGQCVKYDGPKPTHGHWSWSSWSPCSRQYVSHRSLCTNPKPSHGKFCGSTR 630
 QY 668 TLKLCNSQKPRDSVDFRAACACAHNSRFRGRHYKKPYTOVEADLCKLYCIAGDF 727
 DB 631 TLKLCNSQKPRDSVDFRAACACAHNSRFRGRHYKKPYTOVEADLCKLYCIAGDF 690
 QY 728 FSLSNKYVKGTECSEDSRNVCIIGICERVGDNLGSDADEVCGVGNNSACTIHRG 787
 DB 691 FSLSNKYVKGTECSEDSRNVCIIGICERVGDNLGSDADEVCGVGNNSACTIHRG 750
 QY 788 LYTKEHTNOYVHMTTIPSGARSIRIYEMNVSTSVISYRNALRYVYLNGHTVDPGRYK 847
 DB 751 LYTKEHTNOYVHMTTIPSGARSIRIYEMNVSTSVISYRNALRYVYLNGHTVDPGRYK 810
 QY 848 FSGTTFDYRSYNEPNIATGPTNETLIVELLFQGRNPGVAMEYSWRLGTGKOPPAQ 907
 DB 811 FSGTTFDYRSYNEPNIATGPTNETLIVELLFQGRNPGVAMEYSWRLGTGKOPPAQ 870
 QY 908 SYTNIAVRSVCSVSCGGR 926
 DB 871 SYTNIAVRSVCSVSCGGR 899

RESULT 6

AAO15254

ID AAO15254 standard; Protein; 1224 AA.

XX AAO15254;

XX AAO15254;

DT 05-SEP-2002. (first entry)

DE Human MDTs9 protease amino acid sequence.

DE Human; MDTs9 protease; TGF-beta inhibitor;

KW transforming growth factor-beta inhibitor; chronic renal failure.

KW Homo sapiens.

OS WO200251998-A1.

PN 04-JUL-2002.

XX

XX

XX

XX

XX

XX

XX

PF 21-DEC-2001; 2001WO-JP11251.
 PR 25-DEC-2000; 2000JP-0393372.
 PA (YAMA) YAMAMUCHI PHARM CO LTD.
 PI Yamaji N, Nishimura K, Abe K, Ogino M;
 PT WPI; 2002-50888/54.
 DR N-PSDB; AAL43654.
 CC Protease MDTs9 and encoded polynucleotide, applicable in diagnosis and
 screening TGF-beta inhibitors for treatment of chronic renal failure -
 Claim 1; Page 48-51; 60pp; Japanese.
 CC The invention comprises the amino acid and coding sequence of the human
 MDTs9 protease. The MDTs9 DNA and protein sequences of the invention are
 useful for screening for transforming growth factor (TGF)-beta inhibitors
 and for the treatment of chronic renal failure. The present amino acid
 sequence represents the human MDTs9 protease of the invention.
 CC Sequence 1224 AA;
 QY Query Match 73.5%; Score 3846.5; DB 23; Length 1224;
 Best Local Similarity 75.4%; Prod. No. 3e-308;
 Matches 738; Conservative 25; Mismatches 73; Indels 143; Gaps 13;
 QY 1 MKPRARGWGLAALMMLAQVAQVSPGRSHQKNGRSGQLEASPELLSRGPRITAMS 60
 DB 1 MKPRARGWGLAALMMLAQVAQVSPGRSHQKNGRSGQLEASPELLSRGPRITAMS 60
 QY 61 PLSACTCVRHGTRSSAWEPSPASSSTRGAGLCKGRMDVAGNRSQNTYTQENQ 120
 DB 32 PAAA-----PGSSVPRPPPPRPPG-----WHEKG----- 58
 QY 121 TLAVLTQYDAVAYEVDRGDYVSHHMHQRRRAVAVSVESLHLKCPHDPHMDL 180
 DB 59 -----EYDLVAYEVDRGDYVSHHMHQRRRAVAVSVESLHLKCPHDPHMDL 112
 QY 181 RTSSNAPGVITQTLGKTSKVTLPDPFCYQGSLSHNSHSGKFCGSTRTL 240
 DB 113 RTSSNAPGVITQTLGKTSKVTLPDPFCYQGSLSHNSHSGKFCGSTRTL 171
 QY 241 KLCNSQKPRDSVDFRAACACAHNSRFRGR-----HYKKPYTOVEA-DLCKLYCI 291
 DB 172 R-----TEADYFLRLPPLSHLWKL-GRAAQSSPSHVLYKRSTEPHAPGASEVLVT 222
 QY 292 AEGDFPFS--LSNKVKGDTPTCSERDNRVCIIGICELSVSTSAHMPQPKEDFLPDE 349
 DB 223 SRTWELAHQPLHSDLRGLP---QKQHF-----GRRKYMPPQPKEDFLPDE 270
 QY 350 YKSLRHRKSLRSHNEELNVTIYVVDKMMQNHENITTYVLTILNWSALFKD-- 407
 DB 271 YKSLRHRKSLRSHNEELNVTIYVVDKMMQNHENITTYVLTILNWSALFKDGT 330
 QY 408 -----GLMGKDGTRHDHAILLTGLD 427
 DB 331 ICGNINIAVLGLLEDEQPLVISHADHTLSSFCQMSGLMGKDGTRHDHAILLTGLD 390
 QY 428 ICSWKNEPCDTLGFAPISGMSKYRSTINEDTGLGLAFTTAHESGHNFQMHDEGNMC 487
 DB 391 ICSWKNEPCDTLGFAPISGMSKYRSTINEDTGLGLAFTTAHESGHNFQMHDEGNMC 450
 QY 488 KKSEGNIMSPTLAGRNGVFSWSPCSRQYLHKFLSTAQAICLADQPKPKYKPKLPG 547
 DB 451 KKSEGNIMSPTLAGRNGVFSWSPCSRQYLHKFLSTAQAICLADQPKPKYKPKLPG 510
 QY 548 LYDANTQCKWQFGEKAKLMLDKFKDICKALWCHIRGRKCKTFMPAAEGTICGHDWMC 607
 DB 511 LYDANTQCKWQFGEKAKLMLDKFKDICKALWCHIRGRKCKTFMPAAEGTICGHDWMC 570
 QY 608 GGQCVKYDGPKPTHGHWSWSSWSPCSRQYVSHRSLCTNPKPSHGKFCGSTR 667

DB 571 GGCCKVYDGEKPTFHGHWSWSSWPCSRCTCGGVSHRSRLCTNPKPSHGKFCGSGTR 630
 QY 668 TLKLCNSQKCRPSVDFRAAQCAHNSRRFRGRHYKWKPYTQVEDQDLCKLYCIASGPDF 727
 DB 631 TLKLCNSQKCRPSVDFRAAQCAHNSRRFRGRHYKWKPYTQVEDQDLCKLYCIASGPDF 690
 QY 728 FFLSKNKVKDGTCSDESRNVCIDGICERVCNDVLGSDADEVCGVCGNNSACTIHRG 787
 DB 691 FFLSKNKVKDGTCSDESRNVCIDGICERVCNDVLGSDADEVCGVCGNNSACTIHRG 750
 QY 788 LYTGHHTNYQYHMTVTFSGARSIRIYEMNVSTYSIVRNALRYLYLNGHWTYDMPGRYK 847
 DB 751 LYTGHHTNYQYHMTVTFSGARSIRIYEMNVSTYSIVRNALRYLYLNGHWTYDMPGRYK 810
 QY 848 FSGTTFDYRSYNEPENLIATGPTNETLIVELLFQGRNPGVAMEYSMPRLGTEKQPPAOP 907
 DB 811 FSGTTFDYRSYNEPENLIATGPTNETLIVELLFQGRNPGVAMEYSMPRLGTEKQPPAOP 870
 QY 908 SYTMAIVRSECVSCGGGR 926
 DB 871 SYTMAIVRSECVSCGGGR 889

RESULT 7

ABG72431
 ID ABG72431 standard; Protein; 1224 AA.

AC ABG72431;

DT 05-FEB-2003 (first entry)

DE Novel human protease #2.

XW Human; protease; gene therapy; obesity; enzyme.

XX Homo sapiens.

XX US6448388-B1.

FD 10-SEP-2002.

FX 15-AUG-2001; 2000US-0930872.

FX 16-AUG-2000; 2000US-225852P.

XX (LEXI-) LEXICON GENETICS INC.

FX Fiddle C7, Hilbun E;

XX WPI; 2003-074103/07.

DR N-PSDB; ABS57767, ABS57768.

PT New nucleic acid encoding novel human protein (NHP), useful for the preparation of a medicament for diagnosing, preventing or treating disorders or diseases associated with the NHP, e.g., obesity -

XX Claim 4; Column 19-26; 17pp; English.

CC The invention describes a new isolated nucleic acid comprising a sequence that encodes a fully defined protein sequence comprising 491 or 522 amino acids or that hybridises under stringent conditions with the 3675-bp sequence of its complement. The nucleic acid is useful for the preparation of a medicament for diagnosing, preventing or treating disorders (e.g. using gene therapy) of diseases associated with the novel human proteins, e.g., obesity. This sequence encodes a novel human protease.

XX Sequence 1224 AA;

Query Match 73.5%;

Best Local Similarity 75.4%;

Matches 738; Conservative 25; Mismatches 73; Indels 143; Gaps 13;

RESULT 8

AAU80153

ID AAU80153 standard; Protein; 1224 AA.

QY 1 MKFPRARGWGLAALMWLLAQVAQVSPGSHQGRNGSGQLEASPPRLSLRGPRRTAMS 60
 DB 1 MKFPRARGWGLAALMWLLAQVAQVSPGSHQGRNGSGQLEASPPRLSLRGPRRTAMS 31
 QY 61 PLFSGAGTCVRHGTGRSGAWPEPASPASTRTGAAGLOGKGRDMDAENHRSQQINTGTENQ 120
 DB 32 PAAAA-----FGSPSPRPFPABRFG-----VMEKG-----58
 QY 121 TLHLVLTQDVLVSAYEDVRHDGYYSHEIMHQRRRAVAVSEVESLHLRLKGRPHDML 180
 DB 59 -----EYDLVSAYEDVRHDGYYSHEIMHQRRRAVAVSEVESLHLRLKGRPHDML 112
 QY 181 RTSSSLVAPGFIQVTLGKTGTSVQTLPPEDFCFYQGLASHRNSPSHGKFCBGSTRTL 240
 DB 113 RTSSSLVAPGFIQVTLGKTGTSVQTLPPEDFCFYQGLASHRNS-SVALSTCGGLSGMI 171
 QY 241 KLCNSQKCRPSVDFRAAQCAHNSRRFRGR-----HYKWKPYTQVEA-DLCKLYCI 291
 DB 172 R-----TEEDAYFLRPLPSHLNWKD-GRAAQSSPSHVLKYKRSTFEPHAGASVLVT 222
 QY 292 ABGFDFFFS--LSNKVKDGTCSDESRNVCIDGICELSVSTSAHMPQPREOLFLPDE 349
 DB 223 SRTWELAHQPLHSSDLRLGLP--QKQHF-----GRRKKYMPQPREOLFLPDE 270
 QY 350 YKSLRHKRSLRSHRNEBELNVETLVVDKMMQNHENITTYVLTILNMVSALFKO--407
 DB 271 YKSLRHKRSLRSHRNEBELNVETLVVDKMMQNHENITTYVLTILNMVSALFKOGT 330
 QY 408 -----GLMGKDGTRHDHAILLTGLD 427
 DB 331 IGGNINIAIVGLILLEDEQGLVISHHADHTLSSFCQWOSGLMGKDGTRHDHAILLTGLD 390
 QY 428 ICSWKNPECDTGLFAPISGCMKSYRSCITINEDTGLGLAFTIAHESGHNFGIMHDGNNC 487
 DB 391 ICSWKNPECDTGLFAPISGCMKSYRSCITINEDTGLGLAFTIAHESGHNFGIMHDGNNC 450
 QY 488 KXSEGNIMSPTLAGNNGVFSWSPCSRQYLHKFLSTAQAICLADQPKPKYKPEKLGE 547
 DB 451 KXSEGNIMSPTLAGNNGVFSWSPCSRQYLHKFLSTAQAICLADQPKPKYKPEKLGE 510
 QY 548 LYDANTQCKQKQGEKAKLCMLDFKXDKICKALWCHRIKCKCETKFPAAEGTICGHDMMCR 607
 DB 511 LYDANTQCKQKQGEKAKLCMLDFKXDKICKALWCHRIKCKCETKFPAAEGTICGHDMMCR 570
 QY 608 GQCCVKYDGEKPTFHGHWSWSSWSPCSRCTCGGVSHRSRLCTNPKPSHGKFCGSGTR 667
 DB 571 GQCCVKYDGEKPTFHGHWSWSSWSPCSRCTCGGVSHRSRLCTNPKPSHGKFCGSGTR 630
 QY 668 TLKLCNSQKCRPSVDFRAAQCAHNSRRFRGRHYKWKPYTQVEDQDLCKLYCIASGPDF 727
 DB 631 TLKLCNSQKCRPSVDFRAAQCAHNSRRFRGRHYKWKPYTQVEDQDLCKLYCIASGPDF 690
 QY 728 FFLSKNKVKDGTCSDESRNVCIDGICERVCNDVLGSDADEVCGVCGNNSACTIHRG 787
 DB 691 FFLSKNKVKDGTCSDESRNVCIDGICERVCNDVLGSDADEVCGVCGNNSACTIHRG 750
 QY 788 LYTGHHTNYQYHMTVTFSGARSIRIYEMNVSTYSIVRNALRYLYLNGHWTYDMPGRYK 847
 DB 751 LYTGHHTNYQYHMTVTFSGARSIRIYEMNVSTYSIVRNALRYLYLNGHWTYDMPGRYK 810
 QY 848 FSGTTFDYRSYNEPENLIATGPTNETLIVELLFQGRNPGVAMEYSMPRLGTEKQPPAOP 907
 DB 811 FSGTTFDYRSYNEPENLIATGPTNETLIVELLFQGRNPGVAMEYSMPRLGTEKQPPAOP 870
 QY 908 SYTMAIVRSECVSCGGGR 926
 DB 871 SYTMAIVRSECVSCGGGR 889

XX AAU0153; decrease in tumour cell and location of the gene on the 5p-syndrome
 XX 15-JUL-2002 (first entry) CC deletion site on chromosome 5p15.2-15.3 (associated with
 XX Human ADAMTS protein variant. CC Cri-du-chat syndrome). The present sequence represents a variant of
 XX the human ADAMTS protein of the invention.
 XX Sequence 1224 AA;
 XX
 XX Human; ADAMTS; cytostatic; antidiabetic; antithrombotic;
 XX antithrombotic; antitumor; antineoplastic; neovascularisation; angioma;
 XX antidiabetic omentopathy; chronic rheumatoid arthritis; gene therapy;
 XX refractory skin ulcer; gastric ulcer; post-operative healing failure;
 XX repressin-type 2N-metalloproteinase domain; disintegrin-like domain; TSP1;
 XX chromosome 5p15.2-15.3; Cri-du-chat syndrome.
 XX Homo sapiens.
 XX
 XX Key Location/Qualifiers
 XX Misc-difference 104
 XX FT Misc-difference 104 "Wild-type Pro substituted by Ser"
 XX FT Misc-difference 110 "Wild-type Met substituted by Val"
 XX FT Misc-difference 284
 XX FT Misc-difference 284 "Wild-type Ser substituted by Thr"
 XX WO200231163-A1.
 XX
 XX 18-APR-2002.
 XX
 XX 11-OCT-2001; 2001WO-IP08913.
 XX
 XX 11-OCT-2000; 2000JP-0311309.
 XX 02-APR-2001; 2001JP-0102905.
 XX
 XX (KAZU-) KAZUSA DNA RES INST FOUND.
 XX (MITS-) MITSUBISHI PHARMA CORP.
 XX
 XX Ohara O, Nagase T, Nomura N, Yano K, Murakami K, Yasuda S;
 XX Kanazaki K;
 XX
 XX WP1; 2002-372277/40.
 XX
 XX Human brain-originated ADAMTS family polypeptide and encoded gene,
 XX applicable in diagnosis and screening compounds for drug compositions
 XX in treating diseases due to e.g. neovascularisation
 XX
 XX Example 2; Page 1; 172pp; Japanese.
 XX
 XX The invention relates to a polypeptide belonging to the ADAMTS family is
 XX selected from sequences appearing as AAU79496, AAU79497 and AAU79499,
 XX a protein that contains the polypeptide, a protein having not less than
 XX 50% homology with the amino acid sequence of the polypeptides or a
 XX polypeptide modified from any of the polypeptides but with some amino
 XX acids deleted, substituted, added or inserted. Also included are the
 XX polynucleotides encoding the polypeptides (or their complementary
 XX strands or variants), a recombinant vector containing any of the
 XX polynucleotides, a transformant which is transformed with the recombinant
 XX vector, producing the polypeptide, protein or peptide by culturing the
 XX transformant, an antibody that can recognize the polypeptide, protein or
 XX peptide and screening compounds to promote or inhibit activity of the
 XX polypeptide or protein, or to promote or inhibit expression of the
 XX polynucleotide by using the polypeptide, protein, peptide, the
 XX polynucleotide, vector, transformant or/and antibody, particularly in
 XX the presence of a test compound for contact before evaluating the
 XX activity by measuring signal changes. The polypeptide and encoded gene
 XX are applicable in diagnosis and screening compounds for drug compositions
 XX in treating diseases due to neovascularisation, diabetic omentopathy,
 XX chronic rheumatoid arthritis, angioma, refractory skin and gastric ulcers
 XX and post-operative healing failure, including gene therapy.
 XX The gene encoding such polypeptide has conserved repressin-type 2N-
 XX metalloproteinase domain, disintegrin-like domain and TSP1 (thrombospondin
 XX type 1) domain. Its encoded protein is characterised by high expression
 XX in ovaries, changes in expression dose depending on the sexual cycle, a

CC decrease in tumour cell and location of the gene on the 5p-syndrome
 CC deletion site on chromosome 5p15.2-15.3 (associated with
 CC Cri-du-chat syndrome). The present sequence represents a variant of
 CC the human ADAMTS protein of the invention.
 XX Sequence 1224 AA;
 XX
 XX Query Match 73.3%; Score 3840.5; DB 23; Length 1224;
 XX Best Local Similarity 75.3%; Pred. No. 9.5e-308;
 XX Matches 737; Conservative 25; Mismatches 74; Indels 143; Gaps 13;
 QY 1 MKPRARGWGLAALWLLAQAQVAVQSPGRSHQNGRSGQLEASPRLLSRRPRITAMS 60
 DB 1 MKPRARGWGLAALWLLAQAQVAVQSPGRSHQNGRSGQLEASPRLLSRRPRITAMS 60
 QY 61 PLFSAGTCVRHGTGRSGSAWFERPASSTTCAAGLGGKGRMDRAGNHRSSQTNVTENQ 120
 DB 32 PAAAA-----PGSFVRPPPPAERPG-----WMEKG----- 58
 QY 121 TLMVLTQVLYSAYEVDRHGQVYSHEIMHQRRAVAVSEVSLHLKGRPHDPMOL 180
 DB 59 -----EYDLSAYEVDRHGQVYSHEIMHQRRAVAVSEVSLHLKGRPHDPMOL 112
 QY 181 RTSSLVARGFIVQTLGHTGKSNVQLPDPDFCTQSSLSHRNPSHGCGSGTTL 240
 DB 113 RTSSLVARGFIVQTLGHTGKSNVQLPDPDFCTQSSLSHRNPSHGCGSGTTL 171
 QY 241 KLCNSQKCPDUSVDFRAACAEHNSRRFRG-----HYKWKPYTQVEA-DLCKLYCI 291
 DB 172 R-----TEADYFLRPLPSHLSWKL-GRAAQSSPSHVLTKGSTBPAQSEVLVT 222
 QY 282 AEGDFPFFS--LSNKVKDSTPCSDSRNVGIDGICELSVSTSAHWPOPKDELFLDDE 349
 DB 223 SRTWELAQPHSSDURLGLP---QKQHC-----GRKKYMPQPKDELFLDDE 270
 QY 350 YKCSLRHKSLSHRNSELNVELTVVDKMKQNHGHENITTVLTILANVSALFKD-- 407
 DB 271 YKCSLRHKSLSHRNSELNVELTVVDKMKQNHGHENITTVLTILANVSALFKDGT 330
 QY 408 -----GLKXGKGTGRHDAILLTGLD 427
 DB 331 IGGNNIATVGLILLDEPQCLVISHADHTLSSFCQWQSLAKGKGTGRHDAILLTGLD 390
 QY 428 ICSWKNEPCDTLGPAPISGMCVKYSCCTINEDTGLGLAFTTANESCHNFGMIHDEGNWC 487
 DB 391 ICSWKNEPCDTLGPAPISGMCVKYSCCTINEDTGLGLAFTTANESCHNFGMIHDEGNWC 450
 QY 488 KXSEGNINSPTLAAGNVFSPSCSROYLHKFLSTAAQICLAQDPKPKYKPKYKPKLGE 547
 DB 451 KXSEGNINSPTLAAGNVFSPSCSROYLHKFLSTAAQICLAQDPKPKYKPKYKPKLGE 510
 QY 548 LYDANTCKWQFGEKAKLMDLFFKDI CKALWCHIRGKCKETKFMFAEAGTICHHMWR 607
 DB 511 LYDANTCKWQFGEKAKLMDLFFKDI CKALWCHIRGKCKETKFMFAEAGTICHHMWR 570
 QY 608 GGGCVKYGDEGPKPTHGHSWNSWSPSCRTCGGVSRSRLCTNPKPSHGKFCGSGTR 667
 DB 571 GGGCVKYGDEGPKPTHGHSWNSWSPSCRTCGGVSRSRLCTNPKPSHGKFCGSGTR 630
 QY 668 TLKLCNSQKCPDUSVDFRAACAEHNSRRFRGSHYKWKPYTQVEQDLCKLYCIAEGPDF 727
 DB 631 TLKLCNSQKCPDUSVDFRAACAEHNSRRFRGSHYKWKPYTQVEQDLCKLYCIAEGPDF 690
 QY 728 FFSLSNKKVCKDGTCSDESRNVGIDGICERVCNDVLSGDAVEDVCGVGNNSACTIHRG 787
 DB 691 FFSLSNKKVCKDGTCSDESRNVGIDGICERVCNDVLSGDAVEDVCGVGNNSACTIHRG 750
 QY 788 LYTKEHHTNQYHNVITPESGARSIRYEMNVSTSYISVNALRRYVILGHWTVDVDFGKYK 847
 DB 751 LYTKEHHTNQYHNVITPESGARSIRYEMNVSTSYISVNALRRYVILGHWTVDVDFGKYK 810
 QY 848 FSGTTFDVRYSNEPENILATGNTFTLVELLFGQRNFGVAVKESMPRLGTGTEKQPPAQ 907

RESULT 10
 ID ABG76896
 XX ABG76896 standard; Protein; 856 AA.
 XX AC
 XX ABG76896;
 XX DT
 XX 05-NOV-2002 (first entry)
 XX DE Human ADAM-TS 7-like protein #1.
 XX Human; NOVX; cardiomyopathy; atherosclerosis; cell signal processing;
 XX breast cancer; Alzheimer's disease; epilepsy; Huntington's disease;
 XX anxiety; behavioural disorder; multiple sclerosis; myasthenia gravis;
 XX neurodegeneration; Parkinson's disease; pain; stroke; endometriosis;
 XX autoimmune disease; allergy; addition; asthma; transplantation;
 XX graft versus host disease; systemic lupus erythematosus; scleroderma;
 XX psoriasis; Crohn's disease; HIV infection; human immunodeficiency virus;
 XX thrombocytopenia; bleeding disorder; metabolic disorder; obesity;
 XX glucose transport defect; glomerulonephritis; hypercalcaemia;
 XX congenital diarrhoea; respiratory disease; gastro-intestinal disease;
 XX polycystic kidney disease; renal tubular acidosis; skin disorder;
 XX muscle disorder; bone disorder; joint disorder; skeletal disorder;
 XX haematopoietic disorder; urinary system disorder; osteoporosis;
 XX dental disease; dental infection; growth disorder; reproductive disorder;
 XX hypogonadism; fertility disorder; viral infection; bacterial infection;
 XX parasitic infection; metabolic pathway modulation; gene therapy;
 XX zinc metalloprotease; ADAM-TS 7; alpha-2-macroglobulin precursor;
 XX ileal sodium/bile acid cotransporter; prohibitin; MT; CIP4; spinosin;
 XX macrophage stimulating protein precursor; fatty acid-binding protein;
 XX gap junction beta-5 protein; hepsin/plasma transmembrane serine protease.
 XX OS Homo sapiens.
 XX PN W0200233087-A2.
 XX XD 25-APR-2002.
 XX PF 17-OCT-2001; 2001NO-US32496.
 XX PR 17-OCT-2000; 2000US-241040P.
 XX PR 17-OCT-2000; 2000US-241058P.
 XX PR 17-OCT-2000; 2000US-241063P.
 XX PR 17-OCT-2000; 2000US-241243P.
 XX PR 20-OCT-2000; 2000US-242152P.
 XX PR 23-OCT-2000; 2000US-242482P.
 XX PR 23-OCT-2000; 2000US-242611P.
 XX PR 23-OCT-2000; 2000US-242612P.
 XX PR 24-OCT-2000; 2000US-242880P.
 XX PR 24-OCT-2000; 2000US-242881P.
 XX PR 29-DEC-2000; 2000US-259028P.
 XX PR 20-FEB-2001; 2001US-269813P.
 XX PR 25-APR-2001; 2001US-286324P.
 XX PR 29-MAY-2001; 2001US-294108P.
 XX PR 09-JUL-2001; 2001US-303698P.
 XX PR 16-OCT-2001; 2001US-0961151.
 XX PA (CURA-) CURAGEN CORP.
 XX PI Edinger S, Gerlach V, MacDougall JR, Malyankar UM, Smithson G;
 XX PI Millet I, Peyman JA, Stone DJ, Gunther E, Ellerman K, Shinkets RA;
 XX PI Padigaru M, Guo X, Patturajan M, Taupier RJ, Burgess CE;
 XX PI Zehrhusen BD, Kekuda R, Szytek KA, Gangolli EA, Fernandes ER;
 XX PI Gorman L;
 XX DR WPI; 2002-590434/63.
 XX DR N-PSDB; ABS59325.
 XX Cytoplasmic, nuclear, membrane bound and secreted polypeptides and
 XX nucleic acids encoding the polypeptides for diagnosing and treating
 XX e.g. cancer, Alzheimer's disease, cardiomyopathy, metabolic disease and
 XX diabetes.

XX PS
 XX Claim 1; Page 18; 305pp; English.
 XX The present invention relates to new NOVX (NOVI-10) polypeptides. The
 XX molecules of the invention are useful for treating or preventing a
 XX NOVX-associated disorder, such as cardiomyopathy, atherosclerosis, or
 XX a disorder related to cell signal processing and metabolic pathway
 XX modulation in humans. NOVX polypeptides, nucleic acids and antibodies
 XX are useful for treating or preventing disorders or syndromes including
 XX breast cancer, Alzheimer's disease, epilepsy, Huntington's disease,
 XX anxiety, behavioural disorders, multiple sclerosis, myasthenia gravis,
 XX neurodegeneration, Parkinson's disease, pain, stroke, autoimmune
 XX disease, allergies, addition, asthma, endometriosis, graft versus host
 XX disease, systemic lupus erythematosus, scleroderma, transplantation,
 XX psoriasis, Crohn's disease, HIV (human immunodeficiency virus) infection,
 XX atherosclerosis, cirrhosis, rheumatoid arthritis, diabetes,
 XX thrombocytopenia, bleeding disorders, metabolic disorders, obesity,
 XX glucose transport defect, glomerulonephritis, hypercalcaemia, polycystic
 XX kidney disease, pancreatitis, renal tubular acidosis, skin disorders,
 XX congenital diarrhoea, respiratory disease, gastro-intestinal diseases,
 XX muscle, bone, joint and skeletal disorders, haematopoietic disorders,
 XX urinary system disorders, osteoporosis, dental disease and infection,
 XX growth and reproductive disorders, hypogonadism, fertility, and/or other
 XX pathologies and disorders, viral, bacterial, or parasitic infections.
 XX The present amino acid sequence represents a NOVX protein of the
 XX invention.
 XX Sequence 856 AA;
 XX Query Match 64.3%; Score 3365; DB 23; Length 856;
 XX Best Local Similarity 71.9%; Pred. No. 1.1e-268;
 XX Matches 660; Conservative 25; Mismatches 59; Indels 174; Gaps 16;
 QY 129 DLVSAYEVDRHGDYVSHEIMHQRRAVAVSEVSLHLRLKGRHDFHMDLTSSSLVA 189
 DB 19 DLVSAYEVDRHGDYVSHEIMHQRRAVAVSEVSLHLRLKGRHDFHMDLTSSSLVA 75
 QY 189 ---PGFI---VQLGKTKYKSVQLTPED---FCFYQ-GSLRS 221
 DB 76 NSQGFNLINSVSHRQEHASKIITNMLGSSALASTKSNCFVLFSTFFSGNRIKT 135
 QY 222 H-----RNSPSH---GKFCGSG-----TTLKLSQKCFRSDVDFRA 257
 DB 136 EADYPLRPLSHLSKLGRAAQSSPSHVLYKXEVLYTSKTHLAHQ---PLHSDLR 192
 QY 258 AQCAEHNSRFRGHVKNPKYQVEADLKLYCIAGDPFFSLSNKVKDGTCCSDSEN 317
 DB 193 GL---PQKHFCGRKK----- 206
 QY 318 VCIDGICELSVSVTSAAHPPQKEDFLIPDEYKSLRHKSLARHNEELNVLTVV 377
 DB 207 -----YMFQPPREDJFLIPDEYKSLRHKSLARHNEELNVLTVV 250
 QY 378 DKXWQNHGHENITTVTLNMYVSAFKN----- 407
 DB 251 DKXWQNHGHENITTVTLNMYVSAFKN----- 310
 QY 408 -----GLMKDGRHDAHLLTGLDTCWKNEPCDTLGEAPISGKSKYRSCT 455
 DB 311 DHTLSFFCQWQSGLMGKQDTRHDAHLLTGLDTCWKNEPCDTLGEAPISGKSKYRSCT 370
 QY 456 INEDTGLGLAFTTAHSGHNFQIMHDEGNCKKSGNFMSTPLAGRNVFSPSCROY 515
 DB 371 INEDTGLGLAFTTAHSGHNFQIMHDEGNCKKSGNFMSTPLAGRNVFSPSCROY 430
 QY 516 LHKFLSTAQAICLADQPKPKVYKPEKLPGLYDANTOCKWQFGKAKLQWIDFKKDC 575
 DB 431 LHKFLSTAQAICLADQPKPKVYKPEKLPGLYDANTOCKWQFGKAKLQWIDFKKDC 490
 QY 576 VALMCHRIGRKCEKEMPAASGTCGHDMWCRGQCVYGDGPKTHHSDWSSWSPC 635
 DB 491 VALMCHRIGRKCEKEMPAASGTCGHDMWCRGQCVYGDGPKTHHSDWSSWSPC 549

100

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2003, 23:26:49 ; Search time 46 Seconds
(without alignments)
1990.274 Million cell updates/sec

Title: US-09-981-151a-8

Perfect score: 5236

Sequence: 1 MKPARGWRGLAALWMLLAQ.....LEAQQPSATAVIALAFLES 952

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Pir1.*
2: Pir2.*
3: Pir3.*
4: Pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1022.5	19.5	951	2 T00017	Gene ADAMTS-1 prot
2	965.5	18.4	2165	2 T21371	hypothetical prote
3	948	18.1	1205	2 T18157	procollagen N-endo
4	856.5	16.4	837	2 T00355	hypothetical prote
5	731.5	14.0	550	2 T47158	hypothetical prote
6	525.5	10.0	1444	2 T18856	angiogenesis inhib
7	508.5	9.7	1558	2 C89114	protein C37C3.6a
8	508.5	9.7	2167	2 T34395	hypothetical prote
9	499.5	9.5	951	2 T00260	hypothetical prote
10	472.5	9.0	860	2 T16892	hypothetical prote
11	310	5.9	957	2 T15976	hypothetical prote
12	305	5.8	826	2 A60385	monocyte surface a
13	289	5.5	508	2 T22836	hypothetical prote
14	269.5	5.1	616	2 A55796	ecarin precursor
15	263.5	5.0	617	2 T48160	metalloproteinase
16	260	5.0	903	2 S60257	metlatin alpha - mo
17	252.5	4.8	549	2 T48169	metalloproteinase
18	247	4.7	1059	2 T22545	hypothetical prote
19	244.5	4.7	814	2 G02390	disintegrin-like m
20	238	4.5	655	2 JC7850	disintegrin and me
21	236.5	4.5	429	2 A42972	coagulation factor
22	236	4.5	478	2 T34800	fibrinolytic metal
23	232	4.4	1172	1 TSHUP2	thrombospondin 2 p
24	229	4.4	1178	1 A39804	thrombospondin pre
25	228	4.4	1170	1 TSHUP1	thrombospondin 1 p
26	218.5	4.2	1170	2 A40558	thrombospondin 1 p
27	218	4.2	1172	2 A42587	thrombospondin 2 p
28	217	4.1	610	2 JC7530	vascular apoptosis
29	217	4.1	1074	2 JC5928	semaphorin F precu

catrocollastatin p
jatrophin C precu
acetylserine C IEC 3.
acetylserine C IEC 3.
disintegrin-like m
disintegrin (EC 3.
ferritin alpha-II
ferritin alpha-I
metalloproteinase
metalloproteinase
androgen-regulated
metalloproteinase
fibronectin (EC 3.4.
hemorrhagic protei
brain-specific ang
metalloproteinase

ALIGNMENTS

RESULT 1

T00017

Gene ADAMTS-1 protein - mouse

C:Species: Mus musculus (house mouse)

C:Accession: T00017 #sequence_revision 22-Jan-1999 #text_change 21-Jul-2000

E:Kunitz-like domain H.; Ohno, S.; Matsushima, K.

Genomics 46, 466-471, 1997

A:Title: The exon/intron organization and chromosomal mapping of the mouse ADAMTS-1 ge

A:Reference number: 214055; MUID:98110583; PMID:9441751

A:Accession: T00017

A:Notes: Preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-951 <KUN>

A:Cross-references: EMBL:AB001735; NID:G2809056; PIDN:BAA24501.1; PID:G2809057

A:Experimental source: strain 129SVJ

C:Genetics

A:Gene: ADAMTS-1

C:Accession: 228/1; 343/3; 388/1; 444/1; 539/3; 602/1; 660/3; 719/2

C:Superfamily: Chromospondin type 1 repeat homology

F:542-598/Domain: Chromospondin type 1 repeat homology <THR3>

Query Match

Basic Local Similarity 19.5%; Score 1022.5; DB 2; Length 951;

Matches 266; Conservative 99; Mismatches 319; Indels 185; Gaps 32;

QY 167 LELKPRPHDMLTSSLVAPGIVOTLGTGKVSQVLPED---FFCYQGSLSHRH 223

Db 64 LRLDAFQQHLKLPDGGFLGFTQTVGRSGPSEAQLDPTGLAHCFYSGTVNGD- 122

QY 224 NSPSHGKFCGSESTRTKLNCQKCPDSDVDFRAAQAQCAEHNRFRGHYKWKPVQVEA 283

Db 123 -----PGSAALSLCEG-----VRGAFY----- 140

QY 284 DLCKLYCIAGDFFFS-----LSNKVKGDTPCSEDSRN-----VCIDG 322

Db 141 -----LQGEFFIOPAPVATERLAPAVEEESARPOHILRRRRSGGAKC--G 190

QY 323 ICELSVSVTSAMFQ-----PPKEDFLTPDEYKSLRHKRSLASHRNEELN 370

Db 191 VNDDETLFSDRSPESQNTNPNWPRDFTQDAG--KPSGGSIRKRFV-----SSPRY 243

QY 371 VETLVVVQKKNQKHENITTYVLTILNWSALFK----- 406

Db 244 VETLMVADQSMADFHG--SGLKHLLTFVAARFYKHPSTRNSISLVVKVLVIYEEQKG 302

QY 407 -----DGLMGKDGTRHDHAILLTGLDLCISWKNPCCPTLGFAPIS 445

Db 303 PEVTSNAULTLENFNWQKHNSPSDRDPEHYDTAILFTQLDLCG--SHTCDLGMADV 360

QY 446 GNCVKRSTINEDTGLGAFTIAHESGNFMHDEGQNMCKKSEG-----NIMSP 500

```

Db 361 TVCDPSCRSVIEDDGLQAAFTTAHELGHVFNPHD--DAKHCSLNGVTGDSHLMASMLS 419
Qy 501 GRNGVFSWSPSCSRQYLHKFLSTAAQACLAQDPKPKVEKYKPEKLPGLYDANTCKWQRF 560
Db 420 SLDRSOFWSPSCSYWVTSLDNGHGECLMD--KQNPFLKPLSDLPGLTLDANRQCQFTG 477
Qy 561 EKAKLNLDFKKDKALCHRIQRK-----CETKMPAASGTICGHDMWCRGQCQYKG 615
Db 478 BESKHC--PDAACTCTTLC--TGTSGLLVCOTKHPWADGTCGEGKWCYSGKCVNKT 533
Qy 616 D--BQPKPTHGHWDSWSPSCSRCTCGGVSHRSRLCTNPKPSHGKFCGSGSTRTLKLN 673
Db 534 DMKHFATPVHGSWSPGMPAGDCSRTCGGVGYVTRCDNPPVPGKGYCEGRVRYRSCN 593
Qy 674 SOKCP-RDSVDFRAAQAEBHN--SRRPGRH--YKWK-YTQVEDQDLKLYCIAEGFDF 727
Db 594 IEDCPDNNGKTPRECEAHNEFKSFGNEFTVETPKYAGVSPKDRCKLCEAKGIGY 653
Qy 728 FFLSNKVKDGTGPGSEDSNRVCDIGICERVGCDNVLGSDAVEDCVGCGNNSACTIHRG 787
Db 654 FVLQPKVWDGTGPGSPDSTSVCGQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQV 713
Qy 788 LVYTHHTNQYVH-MVTIPSGARSIRYEMV-----STSYISVRNALRYVINGHWTD 841
Db 714 IVTS--TRPGYHOIVTIPAGATNIEVYKRNQGRNRNGSPILARAADGTYLINGNFTLS 770
Qy 842 NPGS-YKFSGTTFYRRSYNEPNIATOPTNETLIVELLFQGR--NPGVAMEXSMRPLG 898
Db 771 TLEQDLTYKGTLYRSGSSAALERIRFSFLKEPLTIQVLMVGHALRPKIKTYPMKK-- 828
Qy 899 TEKOPPAQPSYT-WAIVR-SECVSVCQG 925
Db 829 KTESFNAITFSEWVEEGBECSTCGSG 857

RESULT 2
T21371
hypothetical protein F25H8.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
C:Accession: T21371; T24896
R:Gajadaty, S.
submitted to the EMBL data Library, February 1996
A:Reference number: Z19413
A:Accession: T21371
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2165 <M1>
A:Cross-references: EMBL:269360; PIDN:CA93287.1; GSPDB:GN00022; CESP:F25H8.3
A:Experimental source: clone F25H8
R:Gajadaty, S.
submitted to the EMBL data Library, February 1996
A:Reference number: Z19949
A:Accession: T24996
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2165 <M2>
A:Cross-references: EMBL:269361; PIDN:CA93288.1; GSPDB:GN00022; CESP:F25H8.3
A:Experimental source: clone T13H10
C:Genetics:
A:Gene: CESP:F25H8.3
A:Map position: 4
A:Introns: 31/1; 52/1; 135/2; 193/3; 216/1; 266/1; 495/2; 547/3; 584/3; 634/2; 744/1; 801/1
Query Match 18.4%; Score 965.5; DB 2; Length 2165;
Best Local Similarity 27.9%; Pred. No. 3.4e-61;
Matches 264; Conservative 134; Mismatches 333; Indels 215; Gaps 31;
Qy 122 LHVLTQYDLVSAVEVHGDIYSHVIMHQRARRAVASZ--VESLHLRLKGRPHDFMD 179
Db 66 IHVIDSHHIVRRDSYGRG---KRDVTSRRRLQGVARDGCHACHLRSDRAVIVH 122
Qy 180 LKT-----SSSLVAPGFIVQTKGTGKSVQTLPPDFCFYQGSURS-HR 223

```

RESULT 3

```

T18517
procollagen N-endopeptidase (EC 3.4.24.14) I - bovine
N:Alternate names: procollagen N-proteinase
C:Species: Bos primigenius taurus (cattle)
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T18517
R:Collige, A.; Nuegens, B.V.; Lapiere, C.M.
submitted to the EMBL data Library, February 1996
A:Description: Cloning of the cDNA of the bovine procollagen I N-proteinase.
A:Reference number: T18941
A:Accession: T18517
A:Status: preliminary; translated from GB/EMBL/DBJ

```

```

Db 123 LHRWQKIPDSHNKSPHFSNENFAPMVLYLDSSEEEVFGMSRTDFO--CIVRAHVKGVHQ 180
Qy 224 NSPSHGKFCGSGSTRTLKLNQSK-----CFRSDVDFRAQOC 260
Db 181 HS-----IVNLCSDEGLYGMALPSGIHTVPEIISGNGTTHDG-----ASHR 223
Qy 261 AENNSRRPGRHRYKWKPYQVEADLCKLYCIAEGFDFPFLSNKVKDGTGPGSEDSNRVCI 320
Db 224 RQHLVKRQDPMHFK-----SFQHLNSTSVNFTET----- 253
Qy 321 DGICELSVVSTSAHMQPKEDFLPDBEYKSLRHKRSLRSHRNEELNVELTVVVDKK 380
Db 254 -----VAT-----WQDMEDVIERKARSRAANSWDHYVEVLWADTK 291
Qy 381 MMQNHENITTYLTILNMYSAFKDGLMGKD-----GTR----- 416
Db 292 MYBYHG-ASLEDYVLTFTSVASLYRHQSRLASINVVVYKLVLTENAGRPITQNAQOT 350
Qy 417 -----HDAIILLTGIDICSMKNEPCDTLGFAPISOMCSKYRST 455
Db 351 LQDFCRQYQYNDPDDSSVOHHDVAILTRKDICRSQK--CDTLGLAELGTMDQMKSQA 409
Qy 456 INEDTGLGLAFTIAHESGHNFMIHGE-----GNMCK-----KSEGNIM 495
Db 410 IIEENGLSAAFTIAHELGHVFSIPHODERKCSYMPVNVKVCQKQSTKFDQTFQNNPHIM 469
Qy 496 SPTLAGRNVFSWSPSCSRQYLHKFL--STAQAICLADQPKPKVEKY--YPEKLPGLYDA 551
Db 470 APTLEYNTHPWSWSPSCSAGMLERLENNRGOTQCLFDQPVERRYEDVFRDEPGRKYDA 529
Qy 552 NTQCKWQRFGEAKLMLQPKDICKALMC-----HRIGKCEKTFMPAAEGTIC--CHDM 604
Db 530 HQCKEYFEPASELCPY--MPTCRRLWCATFYGSQWG--CRTQHPWADGTGPGSEDSNR 584
Qy 605 WCRGGQVKYGDGPKPTGHWSWSPSCSRCTCGGVSHRSRLCTNPKPSHGKFCGSG 664
Db 585 FCHGACVRLAPESLTKIDQGMWSWGECSTCGGVQKGLDCDSEKFPNGKYCVG 644
Qy 665 STRTLKLNQSKCPRSDVDFRAAQAEBHNSRRF-----RGRHYKWK-YTQVEDQDLCKL 718
Db 645 QREYRSCNTOECPTQPYREVQCSBEFNKNDIGIGVASINTHWPKYANVAPNERCKL 704
Qy 719 YCIAEGFDFPFLSNKVKDGTGPGSEDSNRVCI DIGICERVGCDNVLGSDAVEDCVGCGN 778
Db 705 YCLRSSAAAPYLLRDVKWDGTPCDRNGDDICVAGACMPAGCDHQLHSTLRDKCGCGSD 764
Qy 779 NSACTHRLGLYTKHHHTNQYVHMVVTIPSGARSIRIYE-----MNVSTSYISVRNALRY 833
Db 765 DSCCKVYKGTNE--QGTFGYENVMKIPAGSANIDIRKQYNNMKEDNDVLSLRANGEP 823
Qy 834 LNHWTVDWP--GRYKFSGTTFYRRSYNEPNIATOPTNETLIVELLFQGRN-PGVAME 891
Db 824 LNHGFQVSLARQQIAFDVLEYSGSDAIIERINGTGPTRSDIYVHVLSVSGSHPPDISYE 883
Qy 892 YSMRL-GTEKQPPAQPSYTAIVR--SECVSVCGGRCCLPVILLE 934
Db 884 YWTAAVENAVIRPISALLYLARVTDVTWTECDRACRQSQSKLMCLD 929

```

A:Molecule type: mRNA
 A:Residues: 1-1205 <COL>
 A:Cross-references: EMBL:X96389; NID:e990769; PID:e228215; PIDN:CA65253.1
 A:Experimental source: skin
 C:Genetics:
 C:Function:
 A:Description: catalyzes cleavage of the propeptides of type I and II collagens prior to
 C:Keywords: hydrolase; metalloproteinase

Query Match 18.1%; Score 948; DB 2; Length 1205;
 Best Local Similarity 29.7%; Pred. No. 3e-60;
 Matches 260; Conservative 131; Mismatches 333; Indels 152; Gaps 30;

QY 137 DHRGVSVSHMHQR-----RRRAVAVESVL-----HLRLGPRDHEM 178
 DB 57 DAQGLVSHVSVAAATAPAGVTRRAAPAIPLSGSEEDPGGLFYVNVVFG-RDLHL 114
 QY 179 DLRSSVAVPGIVOTLTKTKTSVOTLPEDFCYQGLSRHSRHSFGHGFEGSTR 238
 DB 115 RLRFNARLVAPGATVEMQGESGATRV--PILGTCLYGVDAEASSVALSNCDGLAG 172
 QY 239 TLKLNCSQKFRSDVDFRAQCAHNSRRFRG-HYK-KYTOVEADLCKLYCIAEGFD 296
 DB 173 LIRMBEEFFIEPLEKGLAAKEA-----QGRVVVYVHRTTSRPPPL----- 215
 QY 297 FFFSLSNKVDGPPCSSEDSNRVNCIDICELSVVSTSAHMPQPPKEDLFIPLDYKSCURH 356
 DB 216 -----GQFQALDT-GISADSLDLS-----RALGVLEERVNSRRR 250
 QY 357 KRSLRSHRNEELNVETLVVVKMMQNHENITTYVLTILNMYSAFLKDGIMG----- 411
 DB 251 MR---RHAADDYNIENVLLGVDDSVVQFHGTEHVQKYLTLNMIYVHDESLGAHNV 307
 QY 412 -----KDGTF-----RHDHALTLGLDICS 430
 DB 308 VLVRILLSYGKSMSLIEIGNQSLENVCKWAVLQKPDPTDDEVDHDAIFLTRQDF-- 365
 QY 431 WKNPECDTLGAFISGMSKYRSCITNEDTGLGLAFTAHESGHNFGMHGGMWC--K 488
 DB 366 ---GPSGMQYAPVTCMCHPVRSCTLNHEDGFSSAFVVAHETGVLGMEHGDGCGNRGDE 422
 QY 489 KSEGNISPTLAGRNVFSMSPCSQYLLHFLSTAQALCLADQPKVKEYKPEKLPGL 548
 DB 423 VRLGSIAPLVQAAPHRPHMSRCSQOELSYLSYD--CLRDDDF--TDWPAALPQLGLH 479
 QY 549 YDANTQCKWQGEKAKLCLMDPKKIDCKALWCHRIGRK--CETKMPAAEGTICGHDWMC 606
 DB 480 YSNNEQCRFDGFLGYMTCATFRTDFPCQLWCSHPDNPYCTKTKGPELDGTWCAPKHC 539
 QY 607 RGQCVKYQDEGPKPTHGHSWDSWSNSPCSRCTCGGVSHRSRLCTNPKPSHGKCEGST 666
 DB 540 FKGHCIWLTDLIK--RDGNMGAWSPFGSCSRCTCGTGVKTRQCONPHFANGRTCSCLA 598
 QY 667 RTKLKLNCSQKFRSDVDFRAAQA-----EHSNRRFRGRHYKNKPYTOVEDQDLCKLYC 720
 DB 599 YDFQLNCSQCPDALADFBEEQCRQWDLTFEHGDAQ-----HMLPHEHDAKERCHLYC 653
 QY 721 IAEGDFFFSLSNKVKYDGTPCS--EDSRNVICDIGICERVGDNLVGSDAVEDVCGVGN 779
 DB 654 ESKETGEWSMKRMVHDGTSCYSKDAFSLVRCDCRKVCGDGVIGSSQKEDKCGVGDND 713
 QY 780 SACTTHGLYTKHHHTNQYVHMVTPSGARSIRIVEMNVSTYSISVRN-ALRYVYLVNGHW 838
 DB 714 SHCKVVKGTSPKGLYKEMFEIPAGABHLLIQEADTTSHLAVKMLTGTKEFLINEN 773
 QY 839 TVDWPFGRYKFK--SGTTFDYSRVNPEENLIATGPTNETLIVELLFQGRNPGVAMEYS- 894
 DB 774 DVD-FNSKTFIANGVEMEY--ROEDGRETLQTNGLPHGTITVIVIEGQARISLTKYKWIH 831
 QY 895 -PRIGTSKQPPAQP---SYTWAIVR-SECSVSCGG 925
 DB 832 EDSLNVDDNNVLDDSVGYEWALKKMKSPCSKPCGG 867

RESULT 4

100355
 hypothetical protein KIAA0688 - human
 C:Species: Homo sapiens (man)
 C:Date: 01-Feb-1999 Sequence_revision 01-Feb-1999 #text_change 21-Jul-2000
 C:Accession: F00355
 E:Ikeda, K.; Nagase, T.; Suyama, M.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, Y.;
 DNA Res. 5, 169-176, 1998
 A:Title: Prediction of the coding sequences of unidentified human genes. X. The complete
 A:Reference number: Z14142; MUID:98403880; PMID:9734811
 A:Accession: F00355
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-837 <rich>
 A:Cross-references: EMBL:AB014588; NID:g3327189; PIDN:BAA31663.1; PID:g3327190
 A:Experimental source: brain
 C:Genetics:
 C:Function:
 A:Suprafamily: thrombospondin type 1 repeat homology
 F:519-575/Domain: thrombospondin type 1 repeat homology <THR3>

Query Match 16.4%; Score 856.5; DB 2; Length 837;
 Best Local Similarity 35.0%; Pred. No. 7.7e-54;
 Matches 219; Conservative 67; Mismatches 240; Indels 99; Gaps 24;
 QY 371 VETLVVDKMMQNHENITTYVLTILNMYSAFLKDGIMG-----CLMKDGTTRHDH---AIIUTGLDICSMMKNPCDTLGFAPIS 407
 DB 220 VETLVVDKMMQNHENITTYVLTILNMYSAFLKDGIMG-----CLMKDGTTRHDH---AIIUTGLDICSMMKNPCDTLGFAPIS 407
 QY 408 -----CLMKDGTTRHDH---AIIUTGLDICSMMKNPCDTLGFAPIS 445
 DB 279 POGVSAQTLRSFCAMQRLNTPEDSDPDHFTAILFTQDLGVST---CDTLGMADVG 336
 QY 446 GMSKYRSCITNEDTGLGLAFTAHESGHNFGMHGGMWC-----NIMSETL 499
 DB 337 TVCDPAISCAIYEDDGLQSAFAHGLGVFNMHLD--NSKPCISLNGSLSTSHVMAPEVM 395
 QY 500 AGRNVFSMSPCSQYLLHFLSTAQALCLADQPKVKEYKPEKLPGLYDANTQCKWOF 559
 DB 396 AHVDPEEPSPCSARFTDFDLONGVGHCLLD--KPEAPLHLVPTFPQKDYDADRCQLTF 453
 QY 560 GEKAKLCLMDPKKIDCKALWCHRIGRK--KCTKMPAAEGTICGHDWMCRCQCVKYG- 615
 DB 454 GPDNRH--PQLPPCAAUMCSGHLNMGAMCQTRHSPWADGTPCPAQACMGRCRLHMDQ 511
 QY 616 -DEGPKPTHGHSWDSWSNSPCSRCTCGGVSHRSRLCTNPKPSHGKCEGSTRILKLNCS 674
 DB 512 LDQFNIFQAGMGWPGWDCSRCTCGGVQFSRDCRTPVPKNGKQYCEGRTRFRSCNT 571
 QY 675 QKCPRDS--VDFFRAAQAENSR-----RFRGRHYKNKPYTOVEDQDLCKLYCIAEGPDF 727
 DB 572 EDCPTGSALTFRQCAANHRTDLFKSPFG--PMDWYPRYTVAPFDQCKLTCQARALGY 630
 QY 728 FFFSLSNKVDGPPCSSEDSNRVNCIDICERVGDNLVGSDAVEDVCGVGNNSACTIHRG 787
 DB 631 YTVLEPRVVDGTPCPSSESVCGVRCIHAGCDRIIGSKKFKDKWYCGDGSQCSQSG 690
 QY 788 LVTKHHHTNQYVHMVTPSGARSIRIVEMNVSTYSISVRNALR-----RYLLNGHWT---- 839
 DB 691 SFRKFRY--GYNNVVTIPAGATHLVROQG--NPGHRSIYLAUKLPDGSVALNGEYTLMP 747
 QY 840 ---YDWPFG--RYKFSGTTFDYSRVNPEENLIATGPTNETLIVELLFQGRNPGVAMEYS- 893
 DB 748 FTDVDFGAVSLRYSAT-----AASETLSGHLPLAQTFLQVLVAGNPQDTRLKYSF 800
 QY 894 -MPLQTEKQPPAQP---SYTWAIVR-SECSVSCGG 917
 DB 801 FVPR-PTPSTPRTPQ-DWLHRRQA 823

RESULT 5

Db	588	STTLAISEADCPCLSVLQTTTPPMHVVHSDVDFAGKTNPKYKHKHTPFLNENSGMSVWSE	647
Qy	767	AVEDVC-----GVNCG--NNSACT-----IHRGLYTK--HHHTNQYYHMT	803
Db	648	CVTYDCHGTQKVKVRRLCAGALRRBQCTRPCTGSEPLTTFSPQQTFRNFAP	707
Qy	804	IPSGARSIRIYEMNVSTSYISRNALRYLLNGHTVDMFGRYKFGSTTDFYRRSYNEPE	863
Db	708	LP-----NROTNMI-----LRKY--DHWG-PMSAGSVTCGTGQKLRRL--E	743
Qy	864	NLIA-----TGPTNETLIVELLFQGRNPGFVAVESMPRLGTEKPPAQPSYTWAVRSEC	918
Db	744	NCIGQCAETGP-----CVMQSCRENKNTWENS-----QW-----SDC	777
Qy	919	SVSCGGGRCLPVLLEAAC	937
Db	778	SVNCGEG---VQFRKAC	792

C:Species: *Caenorhabditis elegans*
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T16892

RESULT 11

TI15976

hypothetical protein F08C6.1 - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C:Accession: TI15976

R:Sentinel, D.

A:Submitted to the EMBL Data Library, June 1995

A:Description: The sequence of *C. elegans* cosmid F08C6.

A:Reference number: Z18440

A:Accession: TI15976

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule_type: DNA

A:Residues: 1-957 <BEM>

A:Cross-references: EMBL:U9378; NID:G868184; PID:G868185; PIDN:AAA68721.1; CES:12

A:Experimental source: strain Bristol N2

C:Genetics:

A:Gene: CESP:F08C6.1

A:Introns: 23/1; 135/3; 220/3; 285/3; 325/2; 376/3; 461/2; 498/3; 532/3; 627/2

A. Experimental source: strain Bristol N2
A. Genes: C.Gene: CESP:F08C6.1
A. Intron: 23/1; 195/3; 220/3; 285/3; 325/2; 376/3; 461/2; 498/3; 532/3; 627/2;
Query Match 5.9%; Score 310; DB 3; Length 957;
Best Local Similarity 20.8%; Pred. NO. 3.1e-14;
Matches 172; Conservative 73; Mismatches 225; Indels 355; Gaps 41

[illegible]

Db 420 APVTTLPRLSDSHSGGA-TICSOCTNGILSGVGLATARTCSAPYANGSDCVGS 478
 QY 666 TETLKLNSQKCPR--DSVD-PRAAOCABH---NSRPFGRHYHKKPYTOVEDODLCKL 718
 Db 479 TGRAVLC-SRQCGRASVDEYISDCWEQKRLKNDRELTKGSQLNRF-----PQACKV 533
 QY 719 YCIAE-----GPDFFPSLSNKVKGDTFCSDSNVICDGTGCVGCON--VLGSD--A 767
 Db 534 FCDVQHQYSGNRHAF--GNLDPDITSGVD--RYLQDECLAMCNMNAISRDSQC 588
 QY 768 VEDVCGV-----CNG-----NNSACTTHRLTKYKHTHTQYY 799
 Db 589 PDITCPIDTDOSSVVRQGWTSWLTSTCTATCGGYKRNKACSI-TQCCGNEDETEVC 647
 QY 800 HMVTFPGARSIRIYEMNVSTYSIV--RNALRY--YLNCHWTV--DWRGRYVFSG-- 850
 Db 648 SSESCTFSLVAGNENSTWTEWHSVCGRGSGQARYKCLSPHRTLAFDPCGNKYTNEL 707
 QY 851 -TTQDYPRSY--NEPENLTA-----TGPNETLIVELLEFGENGNAWE 891
 Db 708 RITFKARSYIMCVSRCKNTLSKNTSVRSNDGNGFNAI-----GV-WG 754
 QY 892 YSNPRLGTQKQPPAGPSYTWAIVRSESVCSGGKRLPVLLLEAAC 937
 Db 755 -----TWG-GWSTCSTSGFG-----TLVRQRTC 777

RESULT 12

A60385
 monocyte surface antigen MS2 precursor - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 03-Feb-1993 #sequence_revision 03-Feb-1993 #text_change 26-Aug-1999
 C:Accession: A60385
 R:Yoshida, S.; Setauchi, M.; Higuchi, Y.; Akizuki, S.; Yamamoto, S.
 Int. Immunol. 2:585-591, 1990
 A:Title: Molecular cloning of cDNA encoding MS2 antigen, a novel cell surface antigen subunit.
 A:Reference number: A60385, PMID:91197896, PMID:1982220
 A:Accession: A60385
 A:Molecule type: cDNA
 A:Residues: 1-826 <YOS>
 A:Cross-references: EMBL:X13335
 C:Superfamily: mouse mectrin alpha; disintegrin homology
 C:Keywords: glycoprotein; surface antigen; transmembrane protein
 F:14/Domain: signal sequence #status predicted <SIG>
 F:402-484/Domain: disintegrin homology <DIS>
 F:659-683/Domain: transmembrane #status predicted <TM>
 F:330/Active site: Glu #status predicted

Query Match 5.8%; Score 305; DB 2; Length 826;
 Best Local Similarity 21.6%; Pred. No. 5.9e-14;
 Matches 163; Conservative 79; Mismatches 238; Indels 274; Gaps 37;

QY 152 RRRRAVAV---SEVSLHLKGRPHDMDLRTSSSIVAPGFIQTGKTGKSVOTLP 208
 Db 39 RRRRLPSHWGQYPSLSVALGTSGVFTLHLKRDLLGSSVITYSAANGSEVTEQLQ 98
 QY 209 PEDFCFYQGSLSRHSNPSHGKFCBGTSTLLKCNQKC-----PRSDVDFR 256
 Db 99 EQHCLYQGHVEGY-----EGSAASISTCAGLGRFVRGVTHLIEPLDADE-- 145
 QY 257 AAQCAHNSRRFRGRHYKWK-PYTOVEDADLCKLYCIAGDFDFFPSLSNKVKGDTFCSDS 315
 Db 146 -----EGQHAWYQAKHLQKQAGTC-----GVKD----- 168
 QY 316 RNVCDIGICELSVGTSNHPQPKEDLFTLPDEYKSLRHSLSRHSNELAVETLV 375
 Db 169 -----TNLNDLGRALIEYQAPRN--NLIPRETR-----VVELYV 202
 QY 376 VYDKWQNHG--HENITTYVLTILNWSALFKD----- 407
 Db 203 VADSOEFQKLSREAVRQVRVLEWVHVDKLYQLSFRVLVGLIETWKKDYFISYANVT 262
 QY 408 -----GLMKGQTRHDHAILLTGLDICSWNKEPCDTLGFAPISGCMKYSRCTIN 457

Db 263 LENPLSNRQMLQSQH--PHDNVQLITGVDFIG-----STVGLAKVSLCSEH-SGANV 313
 QY 458 ED---TCLGLAFTTASHESHPGMDHGE--GNWC---KKSGNIMSPTLACRNGVFSWS 509
 Db 314 QHNSKNSIGVASTWHELGNLWLSHDEDPICVCPPEGGCGCMTESI-GSKPEPFI 372
 QY 510 PCSQVILHFLSTAAQIADLPK-----PIKEKYKPEKLPDELVDANTOCKWQGEK 562
 Db 373 RSKIDLESFTVKPTGCTGTVNDFVRFVGGFVGNLDFEH--GEQDCQTP----- 422
 QY 563 AKLCHLDFFKOICVALXCHRI-GRKC-----ETKFMFAAB----- 596
 Db 423 -----QCCNQPCNATTCQLVAGRAECASGTCCHECKVKFAGBVCNLSKOKDLEEFCDGR 477
 QY 597 GTICGDMW-----CRGQCQVYGDGPKPTHGNSDSSWSPCSR- 638
 Db 478 KFTCFEDAFQNGTFCPPGSGTFC-----DGSCTFLAQOQR-DLMGQARVAADSCTFFSIPP 532
 QY 639 -CGGCV-SHRSLELTNPKSHGCHGKFGPSTRTLLKNSQKCPRSDVDFRAQCAHNSRR 596
 Db 533 GCGNMTSGRINRC-----GALYCEGQGRPL-----ERSFCT 564
 QY 697 FRGRHYHKKPYTOVEDADLCKLYCIAGDFDFFPSLSNKVKGDTFCSDSNVICDGTG 756
 Db 565 FSSNH-----GVCHALGTGSDITP-----ELVLQTKCEG--KVMDGSCQD 606
 QY 757 VCCDNVLGSDADEVG---GVCGNNSACTIHRG 787
 Db 607 L---RVVRSNCSAKCNHGVGN-HKRECHCKG 636

RESULT 13

T22836
 hypothetical protein F57B7.4 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T22836
 R:Jennings, N.
 Submitted to the EMBL Data Library, June 1996
 A:Reference number: Z19623
 A:Accession: T22836
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-508 <ML>
 A:Cross-references: EMBL:Z74037; PIDN:CAA98493.1; GSPDB:GN00023; CESP:F57B7.4
 C:Genetic code: standard
 C:Experimental source: clone F57B7

Query Match 5.8%; Score 289; DB 2; Length 508;
 Best Local Similarity 23.8%; Pred. No. 4.6e-13;
 Matches 129; Conservative 70; Mismatches 185; Indels 158; Gaps 24;

QY 198 KITGTSVOTLPPEDFCFYQGSLSRHSNPSHGKFCBGTSTRTLLKNSQKCPRSDVDFRA 257
 Db 65 KTQELSDLVLVVADFLSYQAFLEW-SNGDSHR----- 95
 QY 258 AQCCHNSRRFRGRHYKWKPYTOVEDADLCKLYCIAGDFDFFPSLSNKVKGDTFCSDSRN 317
 Db 96 ---AIHN-----LKEYLHALFEQTKI--IYDGISFNWELHWVAGTWATQERD 140
 QY 318 VCIDIGICELSVGTSNHPQPKEDLFTLPDEYKSLRHSLSRHSNELAVETLV 367
 Db 141 -----CPLWI-----SWABEEERVLNREIRLEEKERDLNSTVDDTTFWNSTDS 187
 QY 368 ELAVETLVVD--KQMQ-----NHGHENITTYVLTILNWSALFKDGLMKGQGT--R 416
 Db 189 NSSTDALESMPKLRKEVDFITLEENQENNSTBMTLKIDSKAIDKFTWLKQCTGLP 247
 QY 417 HDHAILLTGLDICSWNKEPCDTLGFAPISGCMKYSRCTINEDTGLGL-AFTIAHESGHN 475


```
QY 580 CH-----RICKCETKEMPAEGTICGHDMKCRGGQCVKYGDBGKPTHGHWSHSPC 635
      |||
      |||
Db 407 CQNFVEVEGED-----CQSTYCR-----NPC 430
      |||
      |||
QY 636 SRTGGGVSHRSLCTNPKPSHGCFEGSTRLKLSQKGRDSVDFRAACQEH--- 692
      |||
      |||
Db 431 -----CNATCKLTP-----GSQADGEC 450
      |||
      |||
QY 693 NSRRGRGHYKPKYTQVEDQDLKLYCIAG-----PDFFSLSNKYVDGTPCSBDSRN 748
      |||
      |||
Db 451 NQCRFRPARTCR--RKIDCDVPE-ICTGSGCEPLDV-----QNGQP-YOSNNGY 500
      |||
      |||
QY 749 CIDGICE--RVGCONV-----LGSDAVEVGVGNGNNS---ACTIHRGLYTKHHHTNQ 797
      |||
      |||
Db 501 CTNGNCEILKNQCIHKNKPAPAGVNVAPDVFEDNQGTQVYNGCIKNGTIYIKAROD- 559
      |||
      |||
QY 798 YVHMVTIPSGARSIIYEMNVSTSIYRNALRYLLNGHWTVDNPGRYKFSGTTFDYR 857
      |||
      |||
Db 560 -----IKCG-----RL 565
      |||
      |||
QY 858 SYNEPENLIATGPTNETLIVELLFOGRNPGVAMEYSMERLQTE 900
      |||
      |||
Db 566 SCEEPS-----TQNTINCAFFSSRP-----DYGWDIGHK 597
      |||
      |||
```

Search completed: October 28, 2003, 23:46:01
Job time : 50 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2003, 21:53:59 ; Search time 48 Seconds
(with alignment)
932.696 Million cell updates/sec

Title: US-09-981-151A-8

Perfect score: 5236

Sequence: 1 MPRARGWGLAALLWLLAQ.....LEACQPSATYALAPLES 952

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2408	46.0	1081	AT18_HUMAN	Q8260 homo sapien
2	1223	23.4	1077	AT10_HUMAN	Q8324 homo sapien
3	1207	23.1	860	AT56_HUMAN	Q8325 homo sapien
4	1205	23.0	1593	AT57_HUMAN	Q8337 homo sapien
5	1189	22.7	997	AT12_HUMAN	Q8344 homo sapien
6	1086	20.8	1906	AT70_MOUSE	Q92541 mus musculus
7	1061	20.3	1629	AT59_HUMAN	Q92541 mus musculus
8	1027	19.6	1911	AT59_HUMAN	Q92541 mus musculus
9	1022	19.5	968	AT51_MOUSE	P97857 mus musculus
10	1021	19.5	967	AT51_MOUSE	P97857 mus musculus
11	1018	19.5	1207	AT19_HUMAN	Q8259 homo sapien
12	998	19.1	1205	AT53_HUMAN	O15072 homo sapien
13	989	18.9	1210	AT19_MOUSE	P59509 mus musculus
14	986	18.8	1095	AT17_HUMAN	Q8256 homo sapien
15	984	18.8	967	AT51_HUMAN	Q8256 homo sapien
16	948	18.1	1205	AT52_BOVIN	P79331 b adamus-2
17	948	18.1	1211	AT52_BOVIN	O95450 h adamus-2
18	929	17.8	950	AT15_HUMAN	Q8258 homo sapien
19	887	16.9	1223	AT14_HUMAN	Q8258 homo sapien
20	886	16.9	930	AT55_MOUSE	Q92001 mus musculus
21	878	16.8	930	AT55_MOUSE	Q92001 mus musculus
22	870	16.6	630	AT34_RAT	Q8267 rattus norvegicus
23	856	16.4	837	AT54_HUMAN	O75173 homo sapien
24	826	15.8	905	AT58_MOUSE	P57110 mus musculus
25	820	15.7	890	AT58_MOUSE	Q5479 homo sapien
26	698	13.3	562	AT15_MOUSE	P59384 mus musculus
27	605	7.7	207	AT55_BOVIN	Q8192 bos taurus
28	340	6.5	245	AT54_BOVIN	Q55492 mus musculus
29	334	6.4	920	AT19_MOUSE	P58459 mus musculus
30	326	6.2	450	AT10_MOUSE	Q5013 homo sapien
31	316	6.0	956	AD08_HUMAN	P78325 homo sapien
32	308	5.9	824	AD08_HUMAN	P78325 homo sapien
33	305	5.8	826	AD08_MOUSE	Q50910 mus musculus

34	281	5.4	857	1	AD22_MOUSE	Q9216 mus musculus
35	269	5.1	616	1	ECAR_ECHCA	Q9495 echis carin
36	260	5.0	819	1	AD09_HUMAN	Q13443 homo sapien
37	260	5.0	903	1	AD12_MOUSE	Q61824 mus musculus
38	257	4.8	697	1	AD26_MOUSE	Q9216 mus musculus
39	255	4.8	739	1	AD26_MOUSE	Q9216 mus musculus
40	252	4.8	739	1	AD26_MOUSE	Q9216 mus musculus
41	245	4.7	813	1	AD35_HUMAN	Q9211 homo sapien
42	243	4.7	813	1	AD35_HUMAN	Q9211 homo sapien
43	238	4.5	1077	1	SW68_MOUSE	Q62217 mus musculus
44	237	4.5	775	1	AD28_HUMAN	Q9432 homo sapien
45	237	4.5	909	1	AD12_HUMAN	Q43184 homo sapien

ALIGNMENTS

RESULT 1					
AT18_HUMAN					
ID	AT18_HUMAN	STANDARD;	PRT:	1081 AA.	
AC	Q8260				
DT	12-SEP-2003 (Rel. 42, Created)				
DT	12-SEP-2003 (Rel. 42, Last sequence update)				
DE	ADAMS-18 precursor (EC 2.14.1.14) disintegrin and metalloproteinase with thrombospondin motifs 13 (ADAM-TS 18) (ADAM-TS18).				
GN	ADAMTS18				
OS	Homo sapiens (Human)				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
FN	[1]				
RP	SEQUENCE FROM N.A.				
RY	MEDLINE=21856482; Pubmed=11867212;				
RA	Cal S., Ohaya A.J., Llamazares M., Garabaya C., Quesada V.,				
RT	"Cloning, expression analysis, and structural characterization of				
RT	seven novel human ADAMTSs, a family of metalloproteinases with				
RT	disintegrin and thrombospondin-1 domains."				
RL	Gene 28349-62(2002);				
CC	-1- COPACOR: Binds 1 zinc ion per subunit (By similarity).				
CC	-1- SUPRACELLULAR LOCATION: Secreted. Associated with the extracellular				
CC	matrix (By similarity).				
CC	-1- TISSUE SPECIFICITY: Expressed in fetal lung, liver, and kidney				
CC	and in adult brain, prostate, submaxillary gland, and endothelium.				
CC	-1- Fm: the precursor is cleaved by a furin endopeptidase (By				
CC	similarity).				
CC	-1- SIMILARITY: Belongs to peptidase family M12B.				
CC	-1- SIMILARITY: Contains 1 disintegrin-like domain.				
CC	-1- SIMILARITY: Contains 4 TSP type-1 domains.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by, and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announcement/				
CC	or send an email to license@isb-sib.ch).				
DR	EMBL; AJ311903; CAC83612.1;				
DR	Genew; HGNC:17110; ADAMTS18.				
DR	MIM; 607512;				
DR	InterPro; IPR002870; Pep_M12B_propep.				
DR	InterPro; IPR001590; Reprolysin.				
DR	InterPro; IPR000884; TSP1.				
DR	Pfam; PF01562; Pep_M12B_propep; 1.				
DR	Pfam; PF01421; Reprolysin; 1.				
DR	Pfam; PF00090; tsp_1; 4.				
DR	SMART; SM00209; TSP1; 4.				
DR	PROSITE; PS50215; ADAM_MEPRO; 1.				
DR	PROSITE; PS00546; CYSTEINE_SWITCH; FALSE NEG.				
DR	PROSITE; PS00427; DISINTEGRIN_1; FALSE NEG.				
DR	PROSITE; PS50214; DISINTEGRIN_2; FALSE NEG.				

PROSITE; PSS0092; TSP1; 3.
 PROSITE; PSS0142; ZINC_PROTEASE; FALSE NEG.
 Repeat; Extracellular matrix.
 SIGNAL
 1 47
 T PROPEP 48 284
 T CHAIN 285 1081
 T DOMAIN 285 497
 T DOMAIN 498 577
 T DOMAIN 588 643
 T DOMAIN 644 878
 T DOMAIN 750 876
 T DOMAIN 877 931
 T DOMAIN 932 946
 T DOMAIN 993 1248
 T SITE 234 235
 T METAL 435 436
 T ACT_SITE 436 436
 T METAL 439 439
 T METAL 445 445
 T CARBOHYD 151 151
 T CARBOHYD 190 190
 T CARBOHYD 313 313
 T CARBOHYD 744 744
 T CARBOHYD 837 837
 T CARBOHYD 908 908
 T SEQUENCE 1081 AA; 0438F6456764616 CRC64;
 Query Match 46.0%; Score 2408; DB 1; Length 1081;
 Best Local Similarity 51.04; Pred. No. 1.7e-167;
 Matches 463; Conservative 118; Mismatches 197; Indels 130; Gaps 18;
 85 ASSTRGAGLGGKGRDMDEAGHRSQQTGTGTQTHLVLTQVPLNSVYDHRGDVUS 144
 47 ADSSSSGAG-----DDYVFVPEVDSAGSTIS 77
 145 HEIMHQRRRAVAVSEVSLHLKPRDFMDLTSSLVAPGFTVQTGKTVKSV 204
 78 HDLHNGKKRS-AQNASSLHRTSFGQLHLK-PSALUSSHFIVQVKGASSET 135
 205 QTLPEDECFVQSLSRHSNPSHGKFCGSGSTRKLKNSQ-----KCPDRSDVFRAQC 260
 136 QK-PEVQCQFVQGRIND-SSSSAVSTAGLGRTRKNEFLISLPQ-----LLA 186
 261 AEHSRRFRGHYKMKPYTQVADLCKLYCIAGDFDFSLSNKVKDG-----TPCSDS 315
 187 QEHYSSPAG-HHPVHLVYKTAEEKIKRGYRGV-----SGRNYGFSPIHSAQS 238
 316 R-----NCVIGICELSVSVSTASAMPQPKEDFLPDPYKSLCRHKSLARS 363
 239 RETEYHRRLOKHPC-----GRKKYAPKPTEDTYLRFDEYSGSGRPRSAKS 289
 364 HRNEELAVETLVVQKQKHQGHENITVYLLTNVSLAFDG----- 408
 290 QKG-LAVETLVVADKREKVGKGNVTYLLVTK-VSLGFKDGTGSDINVVVSLIL 346
 409 -----LMGKDGTRHDHAILLGLDICSWNKSPCDTLGF 441
 347 LEQEPGLLINHHADQSUNSPQMSALIGKNGRKHDAILLTGFDICSWNKSPCDTLGF 406
 442 APTSGMSKYSRCTINEDTGLGLAFTIAHSGHNFQMDHDEGNCKKSGEINFTLIG 501
 407 APTSGMSKYSRCTINEDTGLGLAFTIAHSGHNFQMDHDEGNCKKSGEINFTLIG 466
 502 RNVFWSWSPCSRQYHLKFLSTAQACLADQPKVKEYKYPEKLPGELYDANTCKMORFE 561
 467 NNVFWSWSSCSRQYLLKFLSTPQAGCLVDEPKAQGYKYPKLPQOQIYDADTCKMORFE 526
 562 KAKLWLDFFKDIKALWCHRIKCKETKPNPAEAGTICGHDMWCRGQCCKVYGDGPKP 621
 527 KAKLCSLGFVKDIKSLWCHRVHGRKTPNPAEAGTICGLSMWCRGQCCKVYGDGPKP 586
 632 THGHMSDNSSWSPCSRCTGGGVSHRSLCTNPKFSGHGKFCBGSSTRITLKLCSNCKPCRDS 681

[illegible]

TO A FRAGMENT CONTAINING THE N-TERMINAL REGION INCLUDING THE METALLOPROTEINASE, DISINTEGRIN-LIKE, CYS-RICH AND TS-1 DOMAINS AND THE C-TERMINAL FRAGMENT CONTAINING THE SPACER 2 AND THE FOUR TS-1 DOMAINS.

-/- SIMILARITY: Belongs to peptidase family M12B.

-/- SIMILARITY: Contains 1 disintegrin-like domain.

-/- SIMILARITY: Contains 8 TSP type-1 domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AJ250725; CAC20419.1; -
Genbank; HGNC:14605; ADAMTS12.
MIM; 606184; -
InterPro; IPR001762; Disintegrin.
InterPro; IPR002870; Pep M12B_propep.
InterPro; IPR001590; Reptolysin.
InterPro; IPR000884; TSP1.
InterPro; IPR006025; Zn_M12ptdase.
Pfam; PF01562; Pep M12B_propep.1.
Pfam; PF01421; Reptolysin.1.
Pfam; PF00090; tsep.1; 6.
SMART; SMO0209; TSP1; 8.
PROSITE; PS00215; ADAM_M12PRO; 1.
PROSITE; PS00546; CYSTEINE_SWITCH; FALSE_NEG.
PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
PROSITE; PS00214; DISINTEGRIN_2; FALSE_NEG.
PROSITE; PS00092; TSP1; 6.
ProSite; F800142; ZINC_PROTEASE; 1
Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
Repeat; Extracellular matrix. POTENTIAL
1 25
FT SIGNAL 26 240
FT PROPEP 241 1593
FT CHAIN 241 1593
FT DOMAIN 241 464
FT DOMAIN 465 544
FT DOMAIN 545 577
FT DOMAIN 577 626
FT DOMAIN 626 882
FT DOMAIN 882 942
FT DOMAIN 942 996
FT DOMAIN 996 1315
FT DOMAIN 1312 1365
FT DOMAIN 1367 1421
FT DOMAIN 1422 1470
FT DOMAIN 1471 1531
FT DOMAIN 1531 302
FT DOMAIN 302 305
FT DOMAIN 1538 1570
FT SITE 208 208
FT METAL 392 392
FT ACT SITE 393 393
FT METAL 396 396
FT METAL 402 405
FT METAL 105 105
FT CARBOHYD 125 125
FT CARBOHYD 215 215
FT CARBOHYD 485 485
FT CARBOHYD 685 685
FT CARBOHYD 790 790
FT CARBOHYD 951 951
FT CARBOHYD 1104 1104
FT CARBOHYD 1275 1275
FT CARBOHYD 1300 1300
FT CARBOHYD 1320 1320
FT CARBOHYD 1371 1371
FT CARBOHYD 1371 1378

FT CARBOHYD 1503 1503 N-LINKED (GLCNAC...) (POTENTIAL) .
SQ SEQUENCE 1593 AA; 177455 MW; 07F9F49E63B083A3 CRC64;
Query Match 23.04; Score 1205.5; DB 1; Length 1593;
Best Local Similarity 33.3%; Pred. No. 1e-79;
Matches 288; Conservative 131; Mismatches 311; Indels 135; Gaps 25;
QY 125 LTVQDLVSAYEVDRHGSDYVS---HEINHHQRRRAVAVSEVESLHLRLKGRPHDFHMLR 181
DB 49 LPEYVVGVPRVDASGHPLSYGLHYPTTSRRKRDLDGSE-DWYVYRISHEEKLFNNIT 107
QY 182 TSSSLVAPGFIQVTL-KGTGTFKSVQTLPEPFCFYQGS-LRSHRNSPSHGKFCCESTR 238
DB 108 VNQGFSLNSYINMEKRYGNLSHVGMASAP--LCHLSGTVLQOQTRVGTAAALSACHGLTG 165
QY 239 TLKLCNSQCPKSDVDFRAACAHNSRRFRGH---YMKPKYTOVEADLCKLYCIASG 294
DB 166 FPLQPHG-----DFFTEPVKCHPLVE-GGYHPHIVYRRQKVPETKPTCGL----- 210
QY 295 FDFFFSLSNKVGKGTGTCSDSDSNVCIDGICELSVVSTSAHMPQPPKEDLFLPDEYKSL 354
DB 211 -----KDSVNI-----SQKQELW-----REKWE 228
QY 355 RH---KSLSLASHRNEELAVETLYVVDKMMQNHGHENITVTLNLMWSALFKOGLMG 411
DB 229 RNLPSLSLRSSISKERWVETLYVAUTKMIETVHSGSENVESYILTNMWTGLFHPNSIG 288
QY 412 K-----DGTTHHAILITGL 426
DB 289 NAIHVIVVRLILLEBEOGLKIVHBAKTLSSFCWQKQKINPKSDLPVHRDVAULTK 348
QY 427 DICSWKNECDTLGPAPISGMSKYRSCINEDTGLGLAFTAHESCHNGMIFHDEGNN 486
DB 349 DICAFNRFCETLGLSHLSGCMQPHRSCINEDSLPLATIHLELHGFGLQHDGKND 408
QY 487 CTK--SEGNINSFPLAGNVSFWSPCSQYLHKLSTAAICLAQDPKPVKEVPEKL 544
DB 409 CEPVGRHPIMSRLQVDTPTLWKSCEYITRFDGHWGFLDIPK-KGLSKVIA 467
QY 545 PELDYDANTQKNQFGEKAKLMDLFDKQICAKLWCHRGRKCTKFMFAESTICQDM 604
DB 468 PGVIIVHHQCOLQIGNATPC--QEVENVQCLWKSVMG-FCSRDLDAAGTQCCEKK 524
QY 605 NCRGGQYKYGDGPKPTHUHSWMSGSPGCTGGYSHRSBLCTNPKSHGKCEG 664
DB 525 MCMAGKCIIVYKK-PESIFGGHGRNPSHSCRICGAGQSAERLNNPEPFGKICIG 583
QY 665 STRTLKLSNOKCPDSVDFAACQAHNSRPRGRHYKWKPYTOVEDODLCKLYCIASG 724
DB 584 ERKRYLCNVHCESEAFTRQMCSEFOTVYKNELYHWPFI--FNPAPHCPLYCEPID 641
QY 725 FDFFFSLSNKVGKGTGTCSE--DSRNVICIGICERVCNDVGLSDAIVEDVCGVGNNSAC 782
DB 642 GQFSEKMLDAVIDGTCFEGGNSRNVICINGICWVGCDVEIDSNATEDRCVGLGDSSC 701
QY 783 TTHRGLVTKHHTNOYHYHMTVTPSGARSIRIVENNVSTSYISVENA-LRYYLHGWTVYD 841
DB 702 QTVRKMF-KQKSGSYVDILPKGARDIRVWEIEGAGNFAIRSEDEPKYLANGFIIQ 760
QY 842 WPGRYKFSGTTDVRYSNEPENIATGPTNETLIVELLFQGRNPGVAWEYSMPRLCTEK 901
DB 761 WNGYVKLAVTFQYDRK-GDLEKLMATGPTNESVWQLLFQVTFNGFKIKEYITOKDGLDN 819
QY 902 QPPAQPSYTAIVR-SECSVSCGGG 925
DB 820 D--VEQMYFWQVGHWTCEVTCGTG 842
RESULT 5
ID AT57 HUMAN
AC Q9UKP4; PRT; 997 AA.
DT 16-OCT-2001 (Rel. 40, Created)

DT	16-OCT-2001 (Rel. 40, Last sequence update)
DR	28-FEB-2003 (Rel. 41, Last annotation update)
DE	ADAMS7? precursor [SC 3,4,24], (A disintegrin and metalloproteinase
DE	with thrombospondin motifs 7) (ADAM-TS ?) (ADAM-TS?)
GN	ADAMTS7.
OC	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9506;	[1]
RN	SEQUENCE FROM N.A.
RP	MEDLINE_99395124; PubMed10464288;
RT	Hurskainen T.L., Hirohata S., Seidlin M.F., Apté S.S.:
RT	"ADAM-TS5, ADAM-TS6, and ADAM-TS7, novel members of a new family of
RL	zinc metallopeptidases";
RL	J. Biol. Chem. 274:25555-25563 (1999).
CC	-!- COFACTOR: binds 1 zinc ion per subunit (By similarity).
CC	-!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC	matrix (By similarity).
CC	-!- TISSUE SPECIFICITY: EXPRESSED IN HEART, BRAIN, PLACENTA, LUNG,
CC	LIVER, SKELETAL MUSCLE, KIDNEY, PANCREAS.
CC	-!- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT
CC	FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.
CC	-!- PTM: THE PRECURSOR IS CLEAVED BY A FURIN ENDOPEPTIDASE (BY
CC	SIMILARITY).
CC	-!- SIMILARITY: Belongs to peptidase family M12B.
CC	-!- SIMILARITY: Contains 1 disintegrin-like domain.
CC	-!- SIMILARITY: Contains 1 PLAC domain.
CC	-!- SIMILARITY: Contains 2 TSP type-1 domains.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to licenses@isb-sib.ch).
CC	-----
EMBL	AF140675; RAD56358.1; -.
FR	XESTP_FJ31677; IATL.
DR	NMRP01_M12.223; ADAMTS7.
DR	New; RCSB_PDB: 2Z1.
DR	PMM; Q500G6; ADAMTS7.
DR	InterPro: IPR0011762; Disintegrin.
DR	InterPro: IPR0028970; Rep_M12B_propep.
DR	InterPro: IPR0015900; RPSpolysin.
DR	InterPro: IPR000884; TSPPolysin.
DR	InterPro: IPR0060025; TSP.
DR	InterPro: IPR0060225; Wtpetdse.
DR	pfam; PF01542; Rep_M12B_propep; 1.
DR	pfam; PF01431; RepPolysin; 1.
DR	pfam; PF00930; RPSF1; 2.
DR	SMART; SM00209; TSF1; 2.
DR	PROSITE; PS00142; ZINC_PROTEASE; 1.
DR	PROSITE; PS00215; ADAM_METRO; 1.
DR	PROSITE; PS00546; CYSTEINE_SWITCH; FALSE_NEG.
DR	PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
DR	PROSITE; PS00214; DISINTEGRIN_2; FALSE_NEG.
DR	PROSITE; PSS0092; TSPI; 1.
KW	Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
KW	Repeat; Extracellular matrix.
KW	SIGNAL
FT	CHAIN 1..27 POTENTIAL.
FT	PROPEP 28..232 BY SIMILARITY.
FT	CHAIN 233..997 ADAMTS-7.
FT	DOMAIN 233..461 METALLOPROTEASE.
FT	DOMAIN 462..537 DISINTEGRIN-LIKE.
FT	TSP TYPE-1 538..593 TSP TYPE-1 1.
FT	CYS-RICH 595..697 CYS-RICH.
FT	DOMAIN 698..914 SPACER.
FT	DOMAIN 915..990 TSP TYPE-1 2.
FT	DOMAIN 944..974 PLAC.
FT	SITE 204..204 CYSTEINE SWITCH (POTENTIAL).
FT	METAL 388..388 ZINC (CATALYTIC) (BY SIMILARITY).
FT	ACT SITE 389..389 BY SIMILARITY.

Db 873 RHHLVLMGRPLTQLQESNPGVHYEYTHREAGGHDEYPP--PVFSWHYGPWKCTVT 930

Qy 922 CGGG 925

Db 931 CGRG 934

RESULT 6

AT20 MOUSE
ID AT20 MOUSE STANDARD; PRT; 1906 AA.
AC P59511;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE ADAMS-20 precursor (EC 3.4.24.-) (A disintegrin and
DE metalloproteinase with thrombospondin motifs 20) (ADAM-
DE TS20).
GN ADAMTS20.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ON NCBI_TaxId=10090;
RX [1]
RP SEQUENCE FROM N.A.
RN TISSUE=Petal brain;
RA Llamazares M., Cal S., Lopez-Otin C.;
RT Identification and characterization of ADAMTS-20 defines a novel
RT subfamily of metalloproteinases-disintegrins with multiple
RT thrombospondin-1 repeats and a unique GON-domain.";
RL Submitted (OCT-2002) to the EMBL/GenBank/DDBJ databases.
CC !- FUNCTION: May play a role in tissue-remodeling process occurring
CC in both normal and pathological conditions.
CC !- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC !- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC matrix (By similarity).
CC !- TISSUE SPECIFICITY: Expressed at low level in testis and brain.
CC !- PFM: The precursor is cleaved by a furin endopeptidase (By
CC similarity).
CC !- SIMILARITY: Belongs to peptidase family M12B.
CC !- SIMILARITY: Contains 1 disintegrin-like domain.
CC !- SIMILARITY: Contains 1 GON domain.
CC !- SIMILARITY: Contains 15 TSP type-1 domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AJ512753; CAD54808.3; --
CC DR SMART; SM00209; TSP1, 14.
CC DR PROSITE; PS50215; ADAM_MPRO; 1.
CC DR PROSITE; PS00427; DISINTEGRIN_1; FALSE NEG.
CC DR PROSITE; PS50214; DISINTEGRIN_2; FALSE NEG.
CC DR PROSITE; PS50092; TSP1, 13.
CC DR PROSITE; PS00142; ZINC_PROTEASE; 1.
CC KW Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
KW Repeat; Extracellular matrix.
FT SIGNAL 1 26 POTENTIAL.
FT PROPEP 27 249 BY SIMILARITY.
FT CHAIN 250 1906 ADAMTS-20.
FT DOMAIN 250 464 METALLOPROTEIN-
FT DOMAIN 465 552 DISINTEGRIN-LIKE.
FT DOMAIN 553 608 TSP TYPE-1 1.
FT DOMAIN 609 720 CIS-RICH.
FT DOMAIN 721 842 SPACER.
FT DOMAIN 843 901 TSP TYPE-1 2.
FT DOMAIN 906 962 TSP TYPE-1 3.
FT DOMAIN 962 1015 TSP TYPE-1 4.
FT DOMAIN 1017 1074 TSP TYPE-1 5.
FT DOMAIN 1075 1131 TSP TYPE-1 6.

FT	DOMAIN	1148	1302	TSP TYPE-1 7.
FT	DOMAIN	1200	1360	TSP TYPE-1 8.
FT	DOMAIN	1200	1351	TSP TYPE-1 9.
FT	DOMAIN	1254	1411	TSP TYPE-1 10.
FT	DOMAIN	1412	1465	TSP TYPE-1 11.
FT	DOMAIN	1468	1526	TSP TYPE-1 12.
FT	DOMAIN	1527	1584	TSP TYPE-1 13.
FT	DOMAIN	1585	1648	TSP TYPE-1 14.
FT	DOMAIN	1700	1906	TSP TYPE-1 15.
FT	DOMAIN	1906	1906	GON.
FT	METAL SITE	399	399	ZINC (CATALYTIC) (BY SIMILARITY).
FT	METAL	400	400	ZINC (CATALYTIC) (BY SIMILARITY).
FT	METAL	403	403	ZINC (CATALYTIC) (BY SIMILARITY).
FT	METAL	409	409	ZINC (CATALYTIC) (BY SIMILARITY).
FT	CARBOHYD	52	92	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	721	721	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	721	721	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	798	798	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	805	805	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1057	1057	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1527	1527	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1715	1715	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1759	1759	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1777	1777	N-LINKED (GLCNAC. .) (POTENTIAL).
SQ	SEQUENCE	1906 AA; 212041 MM; 08B256854785578A CRG64;		

Query Match 20.88; Score 1086.5; DB 1; Length 1906;
Best Local Similarity 30.88; Pred. No. 6.1e-71;
Matches 276; Conservative 146; Mismatches 302; Indels 177; Gaps 37;

Qy 125 LTQVDSAYVDHRCGDYVSHMHQRRRAVAVSEVSL--HLKGRPHDFHMDLRT 182

Db 37 LASEVETFTVNEFGD-VFQNHFSRKASGVPEPPFTHRSYAGQLQLNSA 95

Qy 183 SSSIVAPAEFVQLTK----GTGVSQVTLPPEDCFYQSSLSHRNSPSHGKP--CEGS 236

Db 96 DAFLAAGTVEHLGTPEVGGGRSTES-PDLRHCYRQVNAED---HTAVSLCGGL 151

Qy 237 TPTLKLSQKCPDSDVDFA---AQCAEHSRRFRGRHYKKYQTVQADLCKLYCIAE 293

Db 152 MGTEK-ANDGEVLEPV-IRADGSAHDDHN-----KPHLYIQELKR----- 192

Qy 294 GFOFFESLNKVDGTPC-----SEDSNVICIGICLSVSVTSAAHWPOPK 340

Db 193 ---NSFARSHK-----PCVSENMOKTALPQSSRNTGVDVIE-----E 230

Qy 341 EDLFLPDEYKSLRHKRELLSRHNEELN---VETLVVDKQWQKHENITTVLT 396

Db 231 EAVFRLEGE-----RSQLHSNRFLSPRYVEVWNTADAKMHHHG-ONLQHVLT 281

Qy 397 ILANKVSLFKDGLMG-----KDG----- 414

Db 282 LMSIVAIYKQSSIGNLINVIVKLVVHSEGGQVPSFNAATTLNFCMQQQQNVDD 341

Qy 415 ---TRUDHAILGLDICSWNKECTGLGAPISGCKSYKSCITNEDTGLAFTIAHE 471

Db 342 AHPSHDTRAVITREDICQAK-ENCDTGLAELGTLCDEFSKCSISEENGLSAAFTIAHE 400

Qy 472 SCHNFMHDSGNNCK-----KSEGNISPTLAGNNGVSPSCRYLHFLPSTAQAIC 527

Db 401 LGHVFNVPHD--DSFKCRAGTKIQYHVAFTLANTHSPTWASCSOKHTEFLDTGHGEC 459

Qy 528 LAQDPKPYKEYPEKLGELYDANTQCKWQFGEKAKLQMLDFKQIKCAIWC---HRTG 584

Db 460 LLQKPNG-RTDLSPOLSPGVSDGNKQELMFGPGQVC--PYLKH-CHRWLCTSAEGVH 515

Qy 585 PKCTETFWPAEAGTIQGHDMKRGCGQVYQDGGPFPYTHGHSNDSNSPSCRTCCGGVS 644

Db 516 KGCRTQHPPLADGTCGQPMCHRGCLVT-RDMETRPVDGSGWPGWFSYSSCSRTCCGGIK 574

Qy 645 HRSRLCTNPKSHGKFGCEGSTRTLKLNQSKCPROSDVDFRAQCAEHSNRRFRGRHY-- 702

Db 575 STALCDRPEPRNGRYCVGRMKFSCNDCSPKGRGDFRQKQCD-----PDGKHFDI 629

FT VARSPLIC 1073 1529 /FTID=VSP_005499
FT Missing (in isoform Short).
FT /FTID=VSP_005506
FT P -- L (IN REF 1)
SQ SEQUENCE 1629 AA; 182649 MW; C1C4CEFF58B8941F CRC64;

Query Match
Best Local Similarity 27.6%; Pred. No. 3,3e-69;
Matches 200; Conservative 148; Mismatches 334; Indels 253; Gaps 33;

QY 100 RWDDEAGNHRSCQNTGTENGLH-----VLTVDLVSAYVDHRGDVVSHEIMHQ 151
DB 15 RLDAENG-----SPDAAAWEKRLHPQVKLELTGEIYVSPIRVNLGSEFPITNV-HFK 70
QY 152 RRRRAV-----AVSEVSLHLKGPDRPHDMRLTSSIVAPGFIQVTLG 197
DB 71 RTRRSINSATDPWPAFASSSSSTSSQAHYRLSAFGQQLFNLTANAGFIAPLFTVTLG 130
QY 198 KTGKSVOTLPED-----FCFYQSLRSHNSP-----SHGKFKC----- 233
DB 131 TPGVNOTKPYSEAEKLKFCYKVYNTVNTSEHTAVISLCSQMLGTFRSHDGFIFPLQS 190
QY 234 -----EGSTRTLKLNCSQKPR--DSVDFRAACAENSR-----RFRGRHYKKYQV 281
DB 191 MDEQDEERQNPPIIYRSPAPORPESTGRHACDTSEHNHRSHKDKKTRAKRGRINL 250
QY 282 EADLCKLYCIAGDFDFFLSKNVKDGTPGCEDSRNVCDIGELGVSTSAHMPFPKE 341
DB 251 AGDVNALN-----SCLATEFASAYGNKT 273
QY 342 DLFTLPDEYKSLRHSKLSLSHR--NEELAVETLVVDKQKQMGNGHENITTVILTN 399
DB 274 D-----NTEKETHRTKFLSYPRFVVLVADNWSYHG-ENLQHYIILMS 322
QY 400 MVSALFKDGLMK-----DSTRH 417
DB 323 IVASIVKPSIGNLINIVNLVINNEODGPSISFNAQTILNFCQMGHSKSPGLIH 382
QY 418 DIALTLGLDICSNNPECDTLGPAFISQSKYKRCITINETDLGLAFTAHSHNGP 477
DB 383 DVALVLTRODIC-ADKCDLHLAGELGTICOPFRSCSISEDSGLSTAFITHELGVFN 441
QY 478 MHDEGNCK-----ASEGNISYTLAGRVFNSQCSQYHLKFI-STQAQICLADQK 533
DB 442 MHD-UNNKKEGVSPQVFNWFLNFTWPNWMSKSKNITEFLDTGVCGLNLEPE 500
QY 534 PKVYKPKELPGLVDANTQWKGKAKMLQKDKDICALKCHRIG--RVCETK 590
DB 501 S-RYPPLPVQLGLINLVKQCELLFPQSQVCPYMQ--CRELKNVNGVHGKCRTO 556
QY 501 FMPAETGICGDMKRCQCVKGYGDEGPKPTGHNSDASSNSPSCSTCGGVSHSRLC 650
QY 557 HTPWADGTECGSKKYGFCVPKNDVP-VTDGNSGNSPFGTCSRTCGSGIKTAIRC 615
QY 651 TNPESHGKFCBEGSTRLKLNCSQKPRDSVDFRAACAENSRFRGRHY----- 702
DB 616 NRPFGKGYKVGRRMKFKSCNTERCLKRKDRFDEQCA-----HFTGKHFNINGLLPN 670
QY 703 -KWKP-YTOVEDDLCVLTAEAGDFPFSLNKNVKGPCSESRNVICIGICERVGCD 760
DB 671 VNVVPKYSGLMKDKRCKLCFRVAGNTAYQLADRVIDTGPQCGQDNDICVQGLCRQAGCD 730
QY 761 NVLGSDAVEDVCGVGNNSACTHRLGTYKHHHTNYHYHVVTPSGASRIYEMNVS- 819
DB 731 HVLNSKARRDKCGVCGDNSSCKTVAGTNTVHK--CYNTVVRIPAGATINDVQHSFSG 788
QY 820 ----TSYISVRNLRRLYINGHWTVDMPGR-YKFSGTTTIDYRSYENPENLIATGPNET 874
DB 789 ETDDENYALSSKGEFLNGFVVWTAKEIRIRGNNAVVEYSGSETAVERINSFORIEG 848
QY 875 LIVELLPOGR--NPGVANVYSRPLGTGEKQP----- 904
DB 949 LLOVLSVGLKLPNDVRYFENIP---TEDKPQQFYNNHSGPQWQACPKQCGERKRLVCT 905

QY 905 -----AQSRY-----TWAVL-RSECVSVCQ-CGRCLPV 930
DB 906 RESDLQTVSDQRCLRPQGHITPQGTCDLRLHWHVASESCSAQCGLGVRTLDI 960

RESULT 8
AT20_HUMAN STANDARD; PRT; 1911 AA.
AC P59510;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE ADAMTS-20 precursor (EC 3.4.24.-) (A disintegrin and
DE metalloproteinase with thrombospondin motifs 20) (ADAM-TS 20) (ADAM-
TS20).
DS ADAMTS20.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ON NCBI_TaxID:9606;
RX [1]
RN SEQUENCE FROM N.A. (ISOFORM 1).
RP MEDLINE-22511925; PubMed-12514189;
RA Somerville R.P., Longpre J.-M., Jungers K.A., Engle J.M., Ross M.,
RA Evans S., Wright T.N., Leduc R., Apte S.S.;
RA "Characterization of ADAMTS-9 and ADAMTS-20 as a distinct ADAMTS
RA subfamily related to Caenorhabditis elegans GON-1.";
RL J. Biol. Chem. 278:9503-9513(2003).
RN [2]
RN SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RP TISSUE=Liver;
RA Llamazares M., Cal S., Lopez-Otin C.;
RA "Identification and characterization of ADAMTS-20 defines a novel
RA subfamily of metalloproteinases-disintegrins with multiple
RA thrombospondin-1 repeats and a unique GON-domain.";
RT Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: May play a role in tissue remodeling process occurring
CC in both normal and pathological conditions.
CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC matrix (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=P59510-1; Sequence=Displayed;
CC Note=1; P59510-2; Sequence=VSP_007106, VSP_007108,
CC -!- TISSUE SPECIFICITY: Very sparsely expressed although is detected
CC at low levels in testis, prostate, heart, placenta, lung
CC and pancreas. Overexpressed in several brain, colon and breast
CC carcinomas.
CC -!- PM: The precursor is cleaved by a furin endopeptidase (By
CC similarity).
CC -!- SIMILARITY: Belongs to peptidase family M12B.
CC -!- SIMILARITY: Contains 1 disintegrin-like domain.
CC -!- SIMILARITY: Contains 1 GON domain.
CC -!- SIMILARITY: Contains 15 TSP type-1 domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch)
CC EMBL; AF488804; AAC15766.1; --
CC EMBL; AJ515153; CAD56159.3; --
CC EMBL; AY515154; CAD56160.2; --
CC SMART; SM00209; TSP1.12;
CC PROSITE; PS50215; ADAM_MEPRO; 1;
CC PROSITE; PS500427; DISINTEGRIN 1; FALSE NEG

Db 261 VETMLVADQSMADFHG-SGLKYLTLFSAARFYKHPISIRNSISLVVVKILVYBQK 319
 QY 407 -----DLGKDGTRHRAHLLGLDLSKNEPCCTGLGAPLS 445
 Db 320 PEVTSNAULTLNFNCWQKJHNSPSDRPHYDTALLFTQDLG--SHTCOTLGMADV 377
 QY 446 GMSKYRSCINEDTGLGLAFTIAHSGHFGHHDGEGNMCKSEG-----NMSPTLA 500
 Db 378 TVCDPSRSCSVIEDDGLQNAFTAHGLHGVFNPHD-DKHCAASLNGVTGDSHMASLS 436
 QY 501 GRNGVFWSPSCSROVHLKFLSTAOAI CLADQPKPKYKPKLFGELYDMATQCKWQF 560
 Db 437 SLDHQSPWSPCSAYVWTSFLDNGHCEGLND--KPNQPKLPKSDLPGLTDANRQCF 494
 QY 561 EKAKLMLDKPKDICKALCHIRGRK----CETKFPAAAGTICGHDMWCRGQCVKYG 615
 Db 495 ESKKHC--PDAASTCTLLWC--TGTSGGLVLCQTHFFWADGTCGCGKVCVSGKVNKT 550
 QY 616 D--EGPKPTHGHSWMSWSPSCRTCGGVSHRSRLCTNPKPSHGKFCFEGSTRTLKLN 673
 Db 551 DMAKEATFVHSGMGPWGPWDCSRCTCGGVQVYTMRECDNVPKNGKVCYSGKVRYS 610
 QY 674 SQKCP--RDSVDFRAAQAENH--SRFRGRH--YKWKP--YTVQEDQDLCKLYCIAEGFDF 727
 Db 611 IEDCPDNNGKTFRESQCEAHNEFSKASFGNEPTVETPKVAGVSPKDRCKLTCEAKG 670
 QY 728 FSLSNKVKDGTGCSERSNVICIGICERGVCDNVLASDAVEDVCGVCGNMSACTIHRG 787
 Db 671 FFVLOPKVVDGTGCPSTDSVVCQGVKAGCDRIIDSKKPKDCKGVCGGNGSTCKM 730
 QY 788 LYTEKHHNTNOYH--MYTIPSGARSIRYEMNV-----STSYISVRNALRYLNGHMTVD 841
 Db 731 IVTS--TRPGYHDLVTPAGATNLEVHRNQSGRNGNSFLAIRADGDTGYILNGFTLS 787
 QY 842 WPCR--YKPSGTFDYERSYNEPNIAFGPTNETLIVELIFQGR--NPGVAMEYSMPRLG 898
 Db 788 TLEQDLTKYVLYRSGSAAERISRSPEKFTUQVLWGHARLAKIKFTYFMKK-- 845
 QY 899 TEKQPPAPQSYT-WAIVR-SECVSGGG 925
 Db 846 KTESFNAIPFSEWIEWGECSTCGSG 874

RESULT 10
 AT51_RAT1 RAT STANDARD; PRT; 967 AA.
 AC QWUCL, QSERIL, 40, Created
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-FEB-2003 (Rel. 41, Last annotation update)
 DE ADAMTS-1 precursor (EC 3.4.24.-) (a disintegrin and metalloproteinase
 with thrombospondin motifs 1) (ADAM-TS 1) (ADAM-TS1).
 GN ADAMTS
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN NCBI_TaxID=10116;
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Brain;
 RA Liu X., Tu Y., Yin T., Johnstone E.M., Stephenson D.T., Clemens J.A.,
 RA Little S.P.;
 RT Induction of a disintegrin and metalloproteinase with the
 RT thrombospondin type 1 motif (ADAMTS);
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SSSEQUENCE OF 18-967 FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Liver;
 RX MEDLINE=20304099; PubMed=10847486;
 RA Diamantis I., Luetli M., Hoesli M., Reichen J.;
 RT "cloning of the rat ADAMTS-1 gene and its down regulation in
 RT endothelial cells in cirrhotic rats.";
 RL Liver 20:165-172(2000).

CC -- FUNCTION: CLEAVES AGGREGAN, A CARTILAGE PROTEOGLYCAN, AND MAY BE
 CC INVOLVED IN ITS TURNOVER. HAS ANGIOGENIC INHIBITOR ACTIVITY (BY
 CC SIMILARITY). ACTIVE METALLOPROTEASE, WHICH MAY BE ASSOCIATED WITH
 CC VARIOUS INFLAMMATORY PROCESSES AS WELL AS DEVELOPMENT OF CANCER
 CC CACHEKIA. MAY PLAY A CRITICAL ROLE IN FOLLICULAR RUPTURE (BY
 CC SIMILARITY).
 CC -- CATALYTIC ACTIVITY: Cleaves aggrecan at the 1683-Glu-Leu-1684
 CC site, within the chondroitin sulfate attachment domain.
 CC -- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
 CC -- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
 CC matrix (By similarity).
 CC -- INDUCTION: DOWN-REGULATED IN ENDOTHELIAL CELLS DERIVED FROM
 CC CIRRHOTIC LIVER.
 CC -- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT
 CC FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.
 CC -- PTM: THE PRECURSOR IS CLEAVED BY A FURIN ENDOPEPTIDASE (BY
 CC SIMILARITY).
 CC -- SIMILARITY: Belongs to peptidase family M12B.
 CC -- SIMILARITY: Contains 1 disintegrin-like domain.
 CC -- SIMILARITY: Contains 3 TSP type-1 domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation at
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; AF149118; AAC34012.1; -
 CC EMBL; AF304446; AAG29823.1; -
 CC MEROPS; M12.222; -
 CC InterPro; IPR006386; ADAM_cysteine.
 CC InterPro; IPR001762; Disintegrin.
 CC InterPro; IPR002870; Pep_M12B_Prop.
 CC InterPro; IPR001590; Repolysin.
 CC InterPro; IPR000884; TSP1.
 CC InterPro; IPR006025; Zn_MTPetdse.
 CC Pfam; PF01582; Pep_M12B_Prop; 1.
 CC Pfam; PF01421; Repolysin; 1.
 CC Pfam; PF00090; tsp; 1; 3.
 CC SMART; SM00608; ACR; 1.
 CC SMART; SM00209; TSP1; 3.
 CC PROSITE; PS00215; ADAM_MEPRO; 1.
 CC PROSITE; PS00427; DISINTEGRIN; 1.
 CC PROSITE; PS00214; DISINTEGRIN_2; FALSE_NEG.
 CC PROSITE; PS00032; TSP1_PROTEASE; 1.
 CC PROSITE; PS00032; TSP1; 3.
 CC HydroLase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
 CC Signal; Extracellular matrix; Heparin-binding.
 CC FT SIGNAL; 54
 CC FT PROPEP; 55 252
 CC FT CHAIN; 253 967
 CC FT DOMAIN; 253 475
 CC FT DOMAIN; 476 558
 CC FT DOMAIN; 559 614
 CC FT DOMAIN; 616 724
 CC FT DOMAIN; 725 857
 CC FT DOMAIN; 854 910
 CC FT DOMAIN; 911 967
 CC FT DOMAIN; 914 198
 CC FT SITE; 205 205
 CC FT METAL; 401 401
 CC FT ACT SITE; 402 402
 CC FT METAL; 405 405
 CC FT METAL; 411 411
 CC FT CARBOHYD; 547 547
 CC FT CARBOHYD; 720 720
 CC FT CARBOHYD; 764 764
 CC FT CARBOHYD; 782 782
 CC FT CARBOHYD; 945 945
 CC FT CONFLICT; 21 21
 CC FT CONFLICT; 26 31
 CC KFRSSQ -> RSRGSL (IN REF. 2).

862 PENLIATGTTNLT-IVELLFGNRNPOVAWEYS-----MPLRLGTEKOPPAQSYTWAIVR 91

866 MEKLSKAGFTAPLHLPLDLDQNYGLHETITPSDLPENQSSKAP--SPLFWHTS 923

915 -SECSVSCGG 925

924 WEDDTCGG 934

RESULT 12

ATS3 HUMAN STANDARD; PRT; 1205 AA.

16-OCT-2001 (Rel. 40, Created)

28-FEB-2003 (Rel. 41, Last sequence update)

28-FEB-2003 (Rel. 41, Last annotation update)

ADAMTS-3 precursor (SC 3.4.24 (A disintegrin and metalloproteinase with thrombospondin motifs 3) (ADAM-TS 3) (ADAM-TS 3) (Procollagen II (1-NP) aminoproteidase processing enzyme) (Procollagen II N-proteinase) (PC ADAMTS3 OR KIA0366

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

NCBI_TaxID:9606;

[1] SEQUENCE OF 1-227 FROM N.A.

MEDLINE:21402912; PubMed:11409482;

Fernandes R.J, Hirohata S., Engle J.M., Colige A., Cohn D.H., Eyre D.R. Apr 6 S.S:

"Procollagen II amino propeptide processing by ADAMTS-3. Insights on dermatosparaxis."

J. Biol. Chem. 276:31502-31509 (2001).

[2] SEQUENCE OF 5-1205 FROM N.A.

TISSUE:Brain;

MEDLINE:9734984; PubMed:9205841;

Nagase T., Ishikawa K.-I., Nakajima D., Ohira M., Seki N., Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O., "Prediction of the coding sequences of unidentified human genes. VII. The complete sequences of 100 new cDNA clones from brain which can code for large proteins in vitro."

RNA Res. 4:141-150(1997)

1- FUNCTION: Cleaves the propeptides of type II collagen prior to fibril assembly. Does not act on types I and III collagens.

1- COFACTOR: Binds 1 zinc ion per subunit (by similarity).

1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular matrix (by similarity).

1- TISSUE SPECIFICITY: Found in cartilage and skin.

1- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.

1- PTM: THE PRECURSOR IS CLEAVED BY A FURIN ENDOPEPTIDASE (BY SIMILARITY).

1- SIMILARITY: belongs to peptidase family M12B.

1- SIMILARITY: Contains 1 disintegrin-like domain.

1- SIMILARITY: Contains 1 PLAC domain.

1- SIMILARITY: Contains 4 TSP type-1 domains.

1- CAUTION: HAS SOMETIMES BEEN REFERRED TO ADAMTS4.

THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch

EMBL; AF247668; AAK28400.1; --

EMBL; AB002364; BAA20821.1; --

MEROPS; M12.220; --

Genew; HGNC:219; ADAMTS3.

MIM; 605011; --

DR	GO:0005578	Cytoplasmic matrix; NAS.	DR	GO:0005578	Cytoplasmic matrix; NAS.
DR	GO:0004222	F-metallonucleoprotein; NAS.	DR	GO:0004222	F-metallonucleoprotein; NAS.
DR	GO:0004270	F-actin binding activity; NAS.	DR	GO:0004270	F-actin binding activity; NAS.
DR	GO:0003074	P-collagen catabolism; NAS.	DR	GO:0003074	P-collagen catabolism; NAS.
DR	GO:0003099	P-collagen fibril organization; NAS.	DR	GO:0003099	P-collagen fibril organization; NAS.
DR	InterPro:IPR001762	Disintegrin.	DR	InterPro:IPR001762	Disintegrin.
DR	InterPro:IPR002870	Pept M2B protease.	DR	InterPro:IPR002870	Pept M2B protease.
DR	InterPro:IPR001590	Rept M2B protease.	DR	InterPro:IPR001590	Rept M2B protease.
DR	InterPro:IPR000884	TSPI.	DR	InterPro:IPR000884	TSPI.
DR	InterPro:IPR006025	Zn MTpeptidase.	DR	InterPro:IPR006025	Zn MTpeptidase.
DR	Pfam:PF01562	Pept M12B protease; 1.	DR	Pfam:PF01562	Pept M12B protease; 1.
DR	Pfam:PF01421	Reptolysin; 1.	DR	Pfam:PF01421	Reptolysin; 1.
DR	Pfam:PF00090	tsr; 1, 4.	DR	Pfam:PF00090	tsr; 1, 4.
DR	SMART:SM00209	TSPI; 4.	DR	SMART:SM00209	TSPI; 4.
DR	PROSITE:PS02015	ADAM MEFPO; 1.	DR	PROSITE:PS02015	ADAM MEFPO; 1.
DR	PROSITE:PS00427	DISINTEGRIN_1; FALSE NEG.	DR	PROSITE:PS00427	DISINTEGRIN_1; FALSE NEG.
DR	PROSITE:PS02014	DISINTEGRIN_2; FALSE NEG.	DR	PROSITE:PS02014	DISINTEGRIN_2; FALSE NEG.
DR	PROSITE:PS00992	TSPI; 4.	DR	PROSITE:PS00992	TSPI; 4.
DR	PROSITE:PS00142	ZINC PROTEASE; FALSE NEG.	DR	PROSITE:PS00142	ZINC PROTEASE; FALSE NEG.
DR	Hydrolase:Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen; Repeat; Extracellular matrix; Heparin-binding.		DR	Hydrolase:Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen; Repeat; Extracellular matrix; Heparin-binding.	
KW	SIGNAL	1 20	KW	SIGNAL	1 20
FT	PROPEP	21 249	FT	PROPEP	21 249
FT	CHAIN	250 1205	FT	CHAIN	250 1205
FT	DOMAIN	250 469	FT	DOMAIN	250 469
FT	DOMAIN	470 550	FT	DOMAIN	470 550
FT	DOMAIN	551 606	FT	DOMAIN	551 606
FT	DOMAIN	608 712	FT	DOMAIN	608 712
FT	DOMAIN	713 844	FT	DOMAIN	713 844
FT	DOMAIN	845 905	FT	DOMAIN	845 905
FT	DOMAIN	906 965	FT	DOMAIN	906 965
FT	DOMAIN	966 1016	FT	DOMAIN	966 1016
FT	DOMAIN	1020 1050	FT	DOMAIN	1020 1050
FT	DOMAIN	246 249	FT	DOMAIN	246 249
FT	METAL	398 398	FT	METAL	398 398
FT	ACT_SITE	399 399	FT	ACT_SITE	399 399
FT	METAL	402 402	FT	METAL	402 402
FT	METAL	408 408	FT	METAL	408 408
FT	CAREBYND	83 83	FT	CAREBYND	83 83
FT	CAREBYND	119 119	FT	CAREBYND	119 119
FT	CAREBYND	242 242	FT	CAREBYND	242 242
FT	CAREBYND	245 245	FT	CAREBYND	245 245
FT	CAREBYND	813 813	FT	CAREBYND	813 813
FT	CAREBYND	942 942	FT	CAREBYND	942 942
FT	SEQUENCE	1205 AA; EB07B286FC85FB87 CRC64;	FT	SEQUENCE	1205 AA; EB07B286FC85FB87 CRC64;
Query Match	19.18; Score 998; DB 1; Length 1205;		Query Match	19.18; Score 998; DB 1; Length 1205;	
Best Local Similarity	28.68; Pred No. 9,5e-65;		Best Local Similarity	28.68; Pred No. 9,5e-65;	
Matches	260; Conservative 141; Mismatches 355; Indels 152; Gaps 28;		Matches	260; Conservative 141; Mismatches 355; Indels 152; Gaps 28;	
QY	104	EAGHRSQQTNTGENTQTLHLVLTQYDLVSAVEVDHGRDGVSHHEMHORRAVAV-SEV 162	QY	104	EAGHRSQQTNTGENTQTLHLVLTQYDLVSAVEVDHGRDGVSHHEMHORRAVAV-SEV 162
DB	23	QHGNEENVQDLPKIR-----YREYELVPTVSTNLEGRYLSHTLSASHKKSARDYSSNP 77	DB	23	QHGNEENVQDLPKIR-----YREYELVPTVSTNLEGRYLSHTLSASHKKSARDYSSNP 77
QY	163	ESLHLRLKPRDHFMDLTSSSAVAPGVIO-----TLTKGTGKSV 204	QY	163	ESLHLRLKPRDHFMDLTSSSAVAPGVIO-----TLTKGTGKSV 204
DB	78	EQLFFNITAFGDKFHLKLPNTQLVAPGAVVEMHETSILVPGNITDPINNHQSGATVIR 137	DB	78	EQLFFNITAFGDKFHLKLPNTQLVAPGAVVEMHETSILVPGNITDPINNHQSGATVIR 137
QY	205	QTLPEPFCFYGGSLSHRNSPGKFCGSGTTLKLSQKC--PROSVDFRAQCAE 262	QY	205	QTLPEPFCFYGGSLSHRNSPGKFCGSGTTLKLSQKC--PROSVDFRAQCAE 262
DB	138	KTEPLOTNCAVGVDIQDIPCT-SVAISNCGLAGNKSQDVEYPIEPL-----RGKQMBE 192	DB	138	KTEPLOTNCAVGVDIQDIPCT-SVAISNCGLAGNKSQDVEYPIEPL-----RGKQMBE 192
QY	263	HNSRRFRGRH--YKWKPYQVQADLCKLYACBGFDFPFSLSNKKVKGDPGCSDSRNVCI 320	QY	263	HNSRRFRGRH--YKWKPYQVQADLCKLYACBGFDFPFSLSNKKVKGDPGCSDSRNVCI 320
DB	193	EKGR-----HHVYKRSVEQADIDMSK-----DHFRESD-----L 224	DB	193	EKGR-----HHVYKRSVEQADIDMSK-----DHFRESD-----L 224
QY	321	DGICELSVYSTSAHMPQPKEDFILPDBYKSLRKHRSLSFHSNBE-ELNVELVYVVDK 379	QY	321	DGICELSVYSTSAHMPQPKEDFILPDBYKSLRKHRSLSFHSNBE-

CC -!- PWM: The precursor is cleaved by a furin endopeptidase (By
CC similarity).
CC -!- SIMILARITY: Belongs to peptidase family M12B.
CC -!- SIMILARITY: Contains 1 disintegrin-like domain.
CC -!- SIMILARITY: Contains 5 RSP type-1 domains.
CC -!- CAUTION: By homology with the human sequence, it is uncertain
CC whether Met-1 or Met-5 is the initiator.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC FMBI: AV131583; ANM10155.1;
CC WGD; NC1_2442875; Adamts19.
CC SMART; SM00608; ACR; TSP1.5
CC PROSITE; PS00215; ADAM NEPRO; 1
CC PROSITE; PS00546; CYSTEINE SWITCH; FALSE NEG.
CC PROSITE; PS00427; DISINTEGRIN_1; FALSE NEG.
CC PROSITE; PS00264; DISINTEGRIN_2; FALSE NEG.
CC PROSITE; PS00022; TSP; 1
CC PROSITE; PS00022; ZINC_PROTEASE; 1
CC Hydrolase; Metalloproteinase; zinc; Signal; Glycoprotein; Zymogen;
CC Repeat; Extracellular matrix.
CC SIGNAL; 31 319 POTENTIAL
CC CHAIN 31 319 BY SIMILARITY.
CC DOMAIN 320 1210 ADAMTS19
CC DOMAIN 320 1210 METALLOPROTEINASE
CC DOMAIN 320 1210 DISINTEGRIN-LIKE
CC DOMAIN 320 1210 TSP TYPE-1 1.
CC DOMAIN 320 1210 CYS-RICH.
CC DOMAIN 320 1210 SPACER.
CC DOMAIN 320 1210 TSP TYPE-1 2.
CC DOMAIN 320 1210 TSP TYPE-1 3.
CC DOMAIN 320 1210 TSP TYPE-1 4.
CC DOMAIN 320 1210 TSP TYPE-1 5.
CC DOMAIN 320 1210 PLAC.
CC SITE 297 297 CYSTEINE SWITCH (POTENTIAL).
CC METAL 485 485 ZINC (CATALYTIC) (BY SIMILARITY).
CC ACT_SITE 486 486 ZINC (CATALYTIC) (BY SIMILARITY).
CC METAL 489 489 ZINC (CATALYTIC) (BY SIMILARITY).
CC METAL 495 495 ZINC (CATALYTIC) (BY SIMILARITY).
CC CARBOHYD 54 54 N-LINKED GLUCNA. (POTENTIAL).
CC CARBOHYD 263 263 N-LINKED GLUCNA. (POTENTIAL).
CC CARBOHYD 800 800 N-LINKED GLUCNA. (POTENTIAL).
CC CARBOHYD 910 910 N-LINKED GLUCNA. (POTENTIAL).
CC CARBOHYD 931 931 N-LINKED GLUCNA. (POTENTIAL).
CC CARBOHYD 952 952 N-LINKED GLUCNA. (POTENTIAL).
CC CARBOHYD 1012 1012 N-LINKED GLUCNA. (POTENTIAL).
CC SEQUENCE 1210 AA; 134560 MW; 0AB812ABAB7A2 CRC64;
Query Match
Best Local Similarity 28.4%; Pred. No. 4.3e-64;
Matches 295; Conservative 133; Mismatches 362; Indels 250; Gaps 39;

QY 1 MKPRARGWGL-ALNMLLAQVAEVS-----PGRSHORNGRSGOLEAS----- 44
Db 33 LTPDLSEWVFPALMRRESLNATLGSGSDPGSGSGGGRGQASGSRVPSVARA 92
QY 45 PRLSLGPRRLTAMSPLSFAGTCVRHGTSGSAWEPEPASPSTRGACLGKGRWDE 104
Db 93 PQEATRGQSEWFCSP-----EPGADESELESOLPRSSG-----DT 133
QY 105 A---GNRSHQNTYGTENOTLHVLTVDIVSAVEVDHGRGYSHYHMHORRRAVSE 161
Db 134 ALSGGTASWQPLPPORPSSPPAQEPEAEV----- 168
QY 162 VESLHLKLGPRHDFHMDLRTSSSLVAPGFIQTIGKTG---TKSV---QTLPEDFCF 214

Db 169 -----LLRIPALSRDYLILLRRDGRFLAQRFAVEQHPKPGDPTRATADPGSSLLPDASC 224
QY 215 YQGLSRSHNSPESH-----GG-----KFCGSTRITLKLCSKQKPRSDVDFRAA 258
Db 225 YGTVLRHFGSGLASFSTCGGGLMGFIQUNEDFLFIEPFDNTWAIIG----- 270
QY 259 QCAEHNSRFRGRHYKWKPYTOVEADLCKLYCIAEGDFDFFLSNKVKDGTFCSEDSNV 318
Db 271 -----HPRLYRQKR-----STEEKV---TENSVAHRHH 296
QY 319 CIDGICELSVSTSAHMPQPKEDLFLPDEYKSLRHKRSLRSHR-NEELNVTLVVV 377
Db 297 C-----GVISDKGR---PRS-----KKAIDNRKRYSYKLSOEYNIETVVVA 336
QY 378 DKQMOMQNHENITTYVYLILNMVSALFKDGLM----- 410
Db 337 DPAMVSYHGADAARFILTILNMVFNLFQHSLGVQVNLVLKLLIHETPADLVIGHG 396
QY 411 -----GKQGRTHDHAILLTGLDICSMMKNEPCDTLGF 441
Db 397 EKMLSEFCWOHEBFGRNDVHLEMTSMGEDIADVDAAILITRKDFCVHKDEPDTVGI 456
QY 442 APISGMSKYRSCCTINEDTGLGLAFTIAHSGHNFPMIHDSGNMKKSEG-NIMS-PTL 499
Db 457 AYLGMSCKRKCIIAEDNGLNLAFTIAHEMGNHNGINHNDHPSC--ADGLHMSGEMI 514
QY 500 AGRN-GVFSWSPCSROYLHKFLSTAQAICLA-DQPKVKEKYKPEKLFGLDYANTOCKW 557
Db 515 KQNLGDSVMSRCKEDLERFLRSKASSCLLHTDQSLSSVLVPSKLPGMAYTADCCQI 574
QY 558 QFGEKAKLCLMDPKDIOIKALMCHRIQ-KCETKFMPPAAEGTICGHDMWCRGQCVKYGD 616
Db 575 LFGFIASFCQ-EMQHVICTGLMCKVGEAEKCTKLDPPMDGDCDPCGCKACGCTR--- 630
QY 617 EGPXPTH--GHWSWSPSCSRTCCGGVSHRSRLCTNPKPSHGSKFCESGRTTLKLCNS 674
Db 631 RTPAPEHLAGEWSPWS--CSRSCSGVSSSRKRC--FGLSEARDNGPKQVRCEN 685
QY 675 QKCPRSDVDFRAAQCAEHNSRFRGRH-YKNKPYTVQVEDQDLCKLYCIAEGDFDFFLSN 733
Db 686 PPCFAGLPFRDWQCCQYSVTSYPKHALQWQ--AVFDEKPCALFCSPVGKEQPVLLSE 743
QY 734 KYVDGTPCSEDSNRVCIIDGICERVGDVIGSDAIVEDCVGCGNGNSACTIHRGLYTKHH 793
Db 744 KVMGTSYCGYGLDICANGRCQKAGCCGLLSGLAREDHCGVNGNGSKCKVKGDF-NHT 802
QY 794 HTNOYVHMVTIPSGARSIRIYEMVSTYSIVRNALRYLLNGHWTVPWPKYFSGTTP 853
Db 803 RGAGYVEVLVTPAGARRIKVVEEKPAHSLFALRDASK-SINSDKIEHSGAFSLAGTTV 861
QY 854 DYRSYNEPENLIATGPTNETL-IVELLFQGRNFGVAEYSMPRLGTGKEQPPAQ----- 906
Db 862 HYLNR-GLMEKISAKGPTTTLPLHLVLLFPQNYGLHYEYTP---SDPLPNQSSKKEFG 917
QY 907 PSYTWAIVR-SECVSVCGG 925
Db 918 PLFNMWTHAGWDCNATCGG 937
RESULT 14
AT17_HUMAN
ID AT17_HUMAN STANDARD; PRT; 1095 AA.
AC Q8TSE6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE ADAMTS-17 precursor (SC 3.4.24.-) (A disintegrin and metalloproteinase
DE with thrombospondin motifs 17) (ADAM-TS 17) (ADAM-TS17).
GN ADAMTS17.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;

DT 30-MAY-2000 (Rel. 39, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE ADAMTS-1 precursor (SC 3.4.24.-) (A disintegrin and metalloproteinase
 with thrombospondin motifs 1) (ADAM-TS1) (METH-1).
 GN ADAMTS1 OR METH1 OR KIAA1346.
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID:9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Casas C., Pritchard M.A., Estivill X., Arbones M.L.;
 RA "Cloning, characterization and mapping on human chromosome 21 of the
 RA orthologue of murine Adamts-1.";
 RT Submitted (Jul-1999) to the EMBL/GenBank/DBJ databases.
 RL [2]
 RN [3]
 RP SEQUENCE FROM N.A., AND FUNCTION.
 RA TISSUE=Heart;
 RX MEDLINE=99367466; PubMed=10438512;
 RA Vazquez F., Hastings G., Ortega M.-A., Lane T.F., Oikemus S.,
 RA Lombardo M., Inuela-Arigo M.L.;
 RA "METH-1, a human ortholog of ADAMTS-1, and METH-2 are members of a new
 RT family of proteins with angio-inhibitory activity.";
 RN J. Biol. Chem. 274:23349-23357(1999).
 RN [4]
 RP SEQUENCE FROM N.A.
 RA TISSUE=Endothelial cells;
 RX MEDLINE=20247184; PubMed=10785405;
 RA Glienne J., Schmitt A.O., Pilarsky C., Hinzmann B., Weiss B.,
 RA Rosenthal A., Thierach K.H.;
 RA "Differential gene expression by endothelial cells in distinct
 RT angiogenic states.";
 RN Eur. J. Biochem. 267:2820-2830(2000).
 RN [5]
 RP SEQUENCE FROM N.A.
 RA TISSUE=Brain;
 RX MEDLINE=2018126; PubMed=10718198;
 RA Nagase H., Kikuno R., Shirakawa J., Hirose M., Ohara O.;
 RA "Regulation of the coding sequences of unidentified human genes. XVI.
 RT The complete sequences of 15 new cDNA clones from brain which code
 RT for large proteases *in vitro*.";
 RN DNA Res. 7:65-73(2000).
 RN [6]
 RP SEQUENCE FROM N.A.
 RA TISSUE=20289799; PubMed=10830953;
 RA Park H.-S., Fujiwara A., Taylor T.D., Watanabe H., Yada T.,
 RA Soeda E., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Groner Y.,
 RA Polley A., Onki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K.,
 RA Reichwald K., Wenzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
 RA Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,
 RA Shintani A., Sakaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
 RA Minoshima S., Shimizu N., Nordsek G., Hornischer K., Brandt P.,
 RA Schaefer M., Schoen O., Desario A., Reichelt J., Kauer G., Bloeker H.,
 RA Ramser J., Beck A., Klages S., Hennig S., Risselmann L., Daggand E.,
 RA Wehrmeyer S., Bozzym K., Gardiner K., Nizetic D., Francis F.,
 RA Lehmach H., Reinhardt R., Yaspo M.-L.;
 RA "The DNA sequence of human chromosome 21.";
 RN Nature 405:311-319(2000).
 RN [6]
 RP SEQUENCE OF 418-967 FROM N.A.
 RC TISSUE=Melanoma;
 RA Blum H., Baerach S., Mewes H.-W., Well B., Wiemann S.;
 RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- INVOLVED IN LEAVES AGGRECAN, A CARTILAGE PROTEOGLYCAN, AND MAY BE
 CC INVOLVED IN ITS TURNOVER (BY SIMILARITY). HAS ANGIOGENIC INHIBITOR
 CC ACTIVITY. ACTIVE METALLOPROTEASE, WHICH MAY BE ASSOCIATED WITH
 CC VARIOUS INFLAMMATORY PROCESSES AS WELL AS DEVELOPMENT OF CANCER
 CC CACHEXIA. MAY PLAY A CRITICAL ROLE IN FOLLICULAR RUPTURE.
 CC -1- CATALYTIC ACTIVITY: cleaves aggrecan at the 1938-Glu-|-Leu-1939
 CC site, within the chondroitin sulfate attachment domain.
 CC -1- COFACTOR: Binds 1 zinc ion per subunit (by similarity).
 CC

CC -1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
 CC matrix (by similarity).
 CC -1- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT
 CC FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.
 CC -1- PTM: THE PRECURSOR IS CLEAVED BY A FURIN ENDOPEPTIDASE (BY
 CC SIMILARITY).
 CC -1- SIMILARITY: Belongs to peptidase family M12B.
 CC -1- SIMILARITY: Contains 1 disintegrin-like domain.
 CC -1- SIMILARITY: Contains 3 TSP type-1 domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC

DR EMBL: AF170084; AAF15317.1; -;
 DR EMBL: AF060152; AAD48080.1; ALT_INIT.
 DR EMBL: AF207664; AAE23772.1; -;
 DR EMBL: AB037767; BAA92584.1; ALT_INIT.
 DR EMBL: AF001697; BAA95502.1; -;
 DR EMBL: AL162080; CAB82413.1; -;
 DR PIR: T47158; T47158.
 DR MEROPS: M12.222; -;
 DR Genew: HGNC:217; ADAMTS1.
 DR MIM: 605174; -;
 DR GO: GO:0007229; P: integrin-mediated signaling pathway; TAS.
 DR GO: GO:0008285; P: negative regulation of cell proliferation; TAS.
 DR InterPro: IPR006586; ADAM_cysteine.
 DR InterPro: IPR001762; Disintegrin.
 DR InterPro: IPR002870; Pep_M12B_propep.
 DR InterPro: IPR001590; Reprolysin.
 DR InterPro: IPR000884; TSPL.
 DR Pfam: PF01582; Pep_M12B_propep; 1.
 DR Pfam: PF01421; Reprolysin; 1.
 DR Pfam: PF00906; TSR1; 3.
 DR SMART: SM00508; RCR; 3.
 DR PROSITE: PS02015; ADAM_NEPRO; 1.
 DR PROSITE: PS00427; DISINTEGRIN_1; FALSE_NEG.
 DR PROSITE: PS00214; DISINTEGRIN_2; FALSE_NEG.
 DR PROSITE: PS00092; TSPL; 3.
 DR PROSITE: PS00142; ZINC_PROTEASE; 1.
 DR Hydrolase: Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
 RN Repeat; Extracellular matrix; Heparin binding.
 RN SIGNAL 49
 RN PROPEP 50 252
 RN CHAIN 253 967
 RN DOMAIN 253 475
 RN DOMAIN 475 559
 RN DOMAIN 559 614
 RN DOMAIN 617 724
 RN DOMAIN 725 849
 RN DOMAIN 854 905
 RN DOMAIN 908 967
 RN DOMAIN 843 846
 RN SITE 198 198
 RN METAL 401 401
 RN ACT SITE 402 402
 RN METAL 411 411
 RN CARBOHYD 547 547
 RN CARBOHYD 720 720
 RN CARBOHYD 764 764
 RN CONFLICT 227 227
 RN CONFLICT 468 468
 RN CONFLICT 561 561
 RN SEQUENCE 967 AA; 105383 MW; C189389324741ED1 CRC64;
 SQ

Query Match 18.8%; Score 984.5; DB 1; Length 967;

Search completed: October 28, 2003, 23:41:55
Job time : 54 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2003, 23:23:19 ; Search time 76 Seconds
(without alignments)
3232.448 Million cell updates/sec

Title: US-09-981-151A-8

Perfect score: 5236

Sequence: 1 MKPARGRWGLAALWMLLA.....LEAACQPSATYATLAFLES 952

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPRENEL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_micr:*
- 8: sp_micr:*
- 9: sp_micr:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_cheap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3858.5	73.7	1072	4 Q8TE57	Q8TE57 homo sapien
2	3400	64.9	900	11 Q8K206	Q8K206 mus musculus
3	3280	62.6	1021	4 Q8IVE2	Q8IVE2 homo sapien
4	2408	46.0	1081	4 Q8T560	Q8T560 homo sapien
5	2257	43.1	1092	11 Q8BKAL	Q8BKAL mus musculus
6	1218	23.3	1009	11 Q8BKYL	Q8BKYL mus musculus
7	1182	22.6	1070	11 Q8CG28	Q8CG28 mus musculus
8	1075	20.5	1050	5 Q8W493	Q8W493 drosophila
9	1018.5	19.5	1207	4 Q8T559	Q8T559 homo sapien
10	984.5	18.8	967	4 Q8N256	Q8N256 homo sapien
11	965.5	18.4	2165	5 Q19791	Q19791 caenorhabdi
12	944.5	18.0	759	6 Q8ZM48	Q8ZM48 equus caball
13	873.5	16.7	497	11 Q8BZD1	Q8BZD1 mus musculus
14	872.5	16.7	833	11 Q8K384	Q8K384 mus musculus
15	872.5	16.7	845	11 Q8BNJ2	Q8BNJ2 mus musculus
16	861.5	16.5	769	5 Q8MRU5	Q8MRU5 drosophila

17	821.5	15.7	623	11 Q8GSP4	Q8GSP4 mus musculus
18	819.5	15.7	1688	5 Q8SXB0	Q8SXB0 drosophila
19	783	15.0	1427	4 Q8EL37	Q8EL37 homo sapien
20	745	14.2	820	11 Q8C9W3	Q8C9W3 mus musculus
21	640.5	12.2	1235	4 Q8S428	Q8S428 homo sapien
22	632.5	12.1	1280	11 Q8EPX2	Q8EPX2 mus musculus
23	617	11.8	2174	5 Q8GQRO	Q8GQRO drosophila
24	615.5	11.8	3198	5 Q8UG88	Q8UG88 manduca sex
25	615	11.7	3060	5 Q8VAV4	Q8VAV4 drosophila
26	549	10.5	356	4 Q8IWR5	Q8IWR5 homo sapien
27	533	10.2	1091	5 Q8W126	Q8W126 drosophila
28	525.5	10.0	1444	5 Q17591	Q17591 caenorhabdi
29	525.5	10.0	1461	5 Q8WYAB	Q8WYAB caenorhabdi
30	508.5	9.7	1487	5 Q8MPV5	Q8MPV5 caenorhabdi
31	508.5	9.7	1558	5 Q8I710	Q8I710 caenorhabdi
32	508.5	9.7	2167	5 Q876A0	Q876A0 caenorhabdi
33	502	9.6	269	6 Q8GJ54	Q8GJ54 oryctolagus
34	500.5	9.6	790	5 Q8MMEL	Q8MMEL drosophila
35	499.5	9.5	951	4 Q8O345	Q8O345 homo sapien
36	499	9.5	790	5 Q8T458	Q8T458 drosophila
37	489.5	9.3	1572	5 Q44938	Q44938 haemochus
38	473.5	9.0	766	4 P82987	P82987 homo sapien
39	472.5	9.0	872	5 Q22580	Q22580 caenorhabdi
40	470.5	9.0	1055	5 Q19204	Q19204 caenorhabdi
41	461.5	8.8	1020	5 Q8IU50	Q8IU50 caenorhabdi
42	455.5	8.7	439	4 Q8NG66	Q8NG66 homo sapien
43	451.5	8.6	525	4 Q8GRW4	Q8GRW4 homo sapien
44	451.5	8.6	880	5 Q8MSF8	Q8MSF8 drosophila
45	449.5	8.6	377	11 Q8BLI0	Q8BLI0 mus musculus

ALIGNMENTS

RESULT 1

Q8TE57 PRELIMINARY; PRT; 1072 AA.
 AC Q8TE57;
 AT Q8TE57;
 DT 01-JUN-2002 (T-EMBLrel, 21, Created)
 DT 01-JUN-2002 (T-EMBLrel, 21, Last sequence update)
 DT 01-MAR-2003 (T-EMBLrel, 21, Last annotation update)
 DE Metalloprotease disintegrin 16 with thrombospondin type I motif.
 GN ADAMTS16
 OS Homo sapiens (human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21855482; PubMed=11867212;
 RA Cal S., Obaya A.J., Llamazares M., Garabaya C., Quesada V.,
 RA Lopez-Otin C.;
 RT "Cloning, expression analysis, and structural characterization of
 RT seven novel human ADAMTSs, a family of metalloproteinases with
 RT disintegrin and thrombospondin-1 domains";
 RL Gene 283:49-62(2002).
 DR EMBL; AJ315734; CAC86015.1; -;
 DR MEROFS; M12.026; -;
 DR Genew; HGNC:17108; ADAMTS16.
 DR InterPro; IPR006130; Asp/Ora_Cotranf.
 DR InterPro; IPR002870; Pep_M12B_Propesp.
 DR InterPro; IPR001590; Reprolysin.
 DR InterPro; IPR000884; TSP1.
 DR Pfam; PF01562; pep_M12B_Propesp; 1.
 DR Pfam; PF01421; Reprolysin; 1.
 DR Pfam; PF00090; tsp_1; 3.
 DR SMART; SM00209; TSP1; 4.
 DR PROSITE; PS50215; ADAM_MERO; 1.
 DR PROSITE; PS00097; CARMANYLTRANSFERASE; 1.
 DR PROSITE; PS50092; TSP1; 3.
 KW Integrin; Protease.
 SQ SEQUENCE 1072 AA; 119491 MW; 081BEFF78F47D061 CRC64;

Query Match 73.7%; Score 3858.5; DB 4; Length 1072;
 Best Local Similarity 75.64; Pred. No. 0;
 Matches 740; Conservative 24; Mismatches 72; Indels 143; Gaps 13;

QY 1 MKPRABGHWGLAALMLLAQAQVSPCRSHQGRGSGQLASPPRLLSRGPRRLTAMS 60
 DB 1 MKPRABGHWGLAALMLLAQAQVSPCRSHQGRGSGQLASPPRLLSRGPRRLTAMS 60
 DB 1 MKPRABGHWGLAALMLLAQAQVSPCRSHQGRGSGQLASPPRLLSRGPRRLTAMS 60
 QY 61 PLPSAGTCVHRGTSGSNAWPERPASSSTGAGLDCGRDMDEAGNHRSQQTINTG 120
 DB 32 PAAAA-----FGSPSVRPPPPAERFG-----MWKNG-----58
 QY 121 TLHVLTQYDLVSAVYVDRHGDVYSHEIMHQRRAVAVSEVESLHLRLAGPRHMDL 180
 DB 59 -----EYDLVSAVYVDRHGDVYSHEIMHQRRAVAVSEVESLHLRLAGPRHMDL 112
 QY 181 RTSSSLVAPGFIQVTLGKTGTSVOTLPEDFCFYQGLSHRNSPSHGKCEGSTR 240
 DB 113 RTSSSLVAPGFIQVTLGKTGTSVOTLPEDFCFYQGLSHRNSPSHGKCEGSTR 240
 QY 241 KLCNSQKPRDSVDFRAACAEHNSRRPRGR-----HYKWKPYQVEA-DLCLYCI 291
 DB 172 R-----TEEDYFLRPLPULSWKL-GRAAQSGSPSHVLYKSTEPHAPCASEVL 222
 QY 292 AEGDFPFS--LSNWKDGTFCSEDSNRVDCIGICELSVSTVSAHMPQPKDLFLPDE 349
 DB 223 SRTWELAHQPLHSSDLRLGLP--KQHF-----GRRKKYMPQPKDLFLPDE 270
 QY 350 YKSLRHRKSLRSHRNEELNVELVVDKMMQNGHENITTYVLTILNWSALFKD-- 407
 DB 271 YKSLRHRKSLRSHRNEELNVELVVDKMMQNGHENITTYVLTILNWSALFKDGT 330
 QY 408 -----GLMGKDGTRHDHAILLTGLD 427
 DB 331 IGENINIAIVGLILLEDEQGLVISHADHTLSFCQWQGLMGKDGTRHDHAILLTGLD 390
 QY 428 ICNWNKPECTIGFAPISGCMYSKYSTINEDTGLGAPTAHESGNPNFQIHDEGNMC 487
 DB 391 ICNWNKPECTIGFAPISGCMYSKYSTINEDTGLGAPTAHESGNPNFQIHDEGNMC 450
 QY 488 KRSNGINMSTLAGNGVFSNGSCRYLHKLSTQAQICLADQPKVYKYXPKLPG 547
 DB 451 KRSNGINMSTLAGNGVFSNGSCRYLHKLSTQAQICLADQPKVYKYXPKLPG 510
 QY 548 LYDANTCKWQGEKAKLWLDKQICLAKCHRIKCKETKMPAEGTTCGHDMCR 607
 DB 511 LYDANTCKWQGEKAKLWLDKQICLAKCHRIKCKETKMPAEGTTCGHDMCR 570
 QY 608 GGOCVKYDGPRTFHHWDSNNSPCSTCGGVSHRSLCTNPKSHGKPCBASTR 667
 DB 571 GGOCVKYDGPRTFHHWDSNNSPCSTCGGVSHRSLCTNPKSHGKPCBASTR 630
 QY 668 TLKLCNSQKPRDSVDFRAACAEHNSRRPRGRHYKWKPYQVEADQDLCKLYCIAEGDF 727
 DB 631 TLKLCNSQKPRDSVDFRAACAEHNSRRPRGRHYKWKPYQVEADQDLCKLYCIAEGDF 690
 QY 728 FFSLSNKKYDGTFCSEDSNRVDCIGICELSVSTVSAHMPQPKDLFLPDE 787
 DB 691 FFSLSNKKYDGTFCSEDSNRVDCIGICELSVSTVSAHMPQPKDLFLPDE 750
 QY 788 LTKKHHTNQVYHMTVTPSGARSIRIYEMNVSTYSVRNALRYYLNGHWTVDWPGRYK 847
 DB 751 LTKKHHTNQVYHMTVTPSGARSIRIYEMNVSTYSVRNALRYYLNGHWTVDWPGRYK 810
 QY 848 FSGTTFDYRSYNEPENIATGPTNETLIVELLFQGRNPGVAVESMPRLGTEKQPPAQ 907
 DB 811 FSGTTFDYRSYNEPENIATGPTNETLIVELLFQGRNPGVAVESMPRLGTEKQPPAQ 870
 QY 908 SYTWAIVRSECVSCGGGR 926
 DB 871 SYTWAIVRSECVSCGGGG 889

RESULT 2
 QK206 PRELIMINARY; PRT; 900 AA.
 AC QK206
 DT 01-OCT-2002 (TRENBLrel. 22, Created)
 DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
 DE Hypothetical protein.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 NCBI_TaxID=10090;
 RN [1]_TextID=10090;
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC034739; AAH34739.1; -
 DR InterPro; IPR002870; Pep M12B propep.
 DR InterPro; IPR001590; Reprolysin.
 DR InterPro; IPR000884; Tspi.
 DR Pfam; PF01562; Pep M12B propep; 1.
 DR Pfam; PF01421; Reprolysin; 1.
 DR SMART; SM00209; Tspi; 1.
 DR PROSITE; PS00215; ADAM MEPRO; 1.
 DR PROSITE; PS00092; Tspi; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 900 AA; 100679 MW; 590BE2B0E73740FF CRC64;

Query Match 64.9%; Score 3400; DB 11; Length 900;
 Best Local Similarity 65.7%; Pred. No. 3.6e-294;
 Matches 646; Conservative 64; Mismatches 109; Indels 164; Gaps 13;

QY 9 RGLAALW-LLAQAQVSPCRSHQGRGSGQLASPPRLLSRGPRRLTAMSPLFSAGT 67
 DB 4 RGCALWLLLAQAQVSEQTFACALGLAAASGSPEDPQ-----PFF-----45
 QY 68 CVRHGTRSGSNAWPERPASSSTGAGLDCGRDMDEAGNHRSQQTINTGENTLHVLTQ 127
 DB 46 -----SGSSW-----LETF-----55
 QY 128 YDLVSAVYVDRHGDVYSHEIMHQRRAVAVSEV--ESLHLRLAGPRHMDLTS 195
 DB 56 YDLVSAVYVDRHGDVYSHEIMHQRRAVAVTQGGDRLHLKGRPHDHLJDLKAASN 115
 QY 186 LVAPGFIQVTLGKTGTSVOTLPEDFCFYQGLSHRNSPSHGKCEGSTRTLKNS 245
 DB 116 LVAPGFIQVTLGKTGTSVOTLPEDFCFYQGLSHRNSPSHGKCEGSTRTLKNS 170
 QY 246 QKPRDSVDFRAACAEHNSRR-----GRYKWKPYQVEA-----DL 285
 DB 171 -----TKDIDYFLPPLPULSWKLNSAQGSPPHLYRSTEQAPRENVLMITRKDL 226
 QY 286 CKLYCIAEGDFPFSLSNKKYDGTFCSEDSNRVDCIGICELSVSTVSAHMPQPKDLFI 345
 DB 227 ARPLHNDHNLGSPKQHF-----CGRKK-----YMFQPPDDLYI 264
 QY 346 LPDEYKSCLPKHSLLRSHRNEELNVELVVDKMMQNGHENITTYVLTILNWSALF 405
 DB 265 LPDEYKSPSSRHSLSLHSHRNEELNVELVVDKMMQNGHENITTYVLTILNWSALF 324
 QY 406 KD-----GLMGKDGTRHDHAILLT 423
 DB 325 KDTIGGININIVGLILLEDEQGLAISHDHAILTTSFCQWQSGLCKGKGTTHDAILL 384
 QY 424 TGLDICSWKNEPCDTIGFAPISGCMYSKYSTINEDTGLGAPTAHESGNPNFQIHDE 483
 DB 395 TGLDICSWKNEPCDTIGFAPISGCMYSKYSTINEDTGLGAPTAHESGNPNFQIHDE 444
 QY 484 GNMCKKSEGINMSTLAGNGVFSNGSCRYLHKLSTQAQICLADQPKVYKYXPK 543
 DB 445 GNMCKKSEGINMSTLAGNGVFSNGSCRYLHKLSTQAQICLADQPKVYKYXPK 504

QY 412 -----XGTRHDHALLTGLDLSKWNKPCDGLGAPISQNSCKYBSCTINEDYGLGLAF 466
 DB 413 NKSGEKDPHHHOVAILLTKKICA---NNCVTLGLANVGWCKPQKQSVNEDGIMLSH 469
 QY 467 TIAHSGHNFHMGHDEGCMCKSEG---NIMSTTLAQRNVSFMSPCSRDYLHKLSTA 523
 DB 470 TITHELGHNFGRHDTAKIGCHPRVGPVILHMTFTGADTLQVCMSNCRKXITIFLQGG 529
 QY 524 CAICLADQPKVPKVEYKPEKLPGLDYDANTCKWQFG--EKAKLCLMDPKKIDICALWCH 581
 DB 530 LGERCLDDPPTPLDEYNYVTGELFGMRNARGOCLQFNLTITDSEVGACSAPEHFCSTLWC- 588
 QY 582 RIGSKCTETKFWPAAETICGHDMMWCGGQCVKYGDEGPRKPTHGHSWSSWSPSCSTCGG 641
 DB 589 KVNCEGCVTHMFTPTAGTLCGNKWCQNGKVR--REELAAVNGGWDGMSWSECSRSQGG 646
 QY 642 GVSHRRLCTNPKSHGGKFCGEGSTRTKLKNSQKCPDROSVDFRAAQCAEHNSRRFRGRH 701
 DB 647 GVSTQRECDNFVANGVFCIGERRKYKICRKPCEPABEESFRAQCCARFQNVSTQAT 706
 QY 702 YAKPYTVOEDQDLKLYCIAEGDFFFSLSNKYKDGTPCSEDSRVNCDIGICERYGCDN 761
 DB 707 YKMLFP--PDKNPKLFCSDVDPTIANWGAATVLDGTPCTLNNMCDIGICKYVGDW 764
 QY 762 VLGSDAVEDVCGVGNNSACTIHRGLITKHHHTNQ--YHHMTIPSGARSIRIYEMNVS 819
 DB 765 IVDSVQDQKCGVSGSGQCPVRETYTDFPAKAGAYBELVTFARAHILIRLANS 824
 QY 820 TSYISVR--NALRYLNGHVTVDGVRKFSCTTFDVRSENPENLIATGPTNETHIV 877
 DB 825 PHFIATAGDGGDFYNGDSLIMFGEFAGAESLYDR-VDEQSTIIPQIQRSHSL 883
 QY 878 ELLPQR--NPGVAWESYBPLGTEKQPAQPSYVAIVR--SEGSUSGGRCGLPVLLE 934
 DB 884 YAIVRGNSAGIYFETPLALNV---TAGRFQWLSNWTACASACGG---VQRE 935
 QY 935 AACQ 938
 DB 936 PTCQ 939

RESULT 9
 ID Q8TE59 PRELIMINARY; PR3; 1207 AA.
 AC Q8TE59;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE ADAMS-19.
 GN ADAMS19.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID:9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21856482; PubMed=11867212;
 RA Cal S., Ojaya A.J., Llamazares M., Garabaya C., Quesada V.,
 RA Lopez-Otin C.;
 RT "Cloning, expression analysis, and structural characterization of
 RT seven novel human ADAMS19, a family of metalloproteinases with
 RT disincergin and thrombospondin-1 domains.";
 RL Gene 283:49-62(2002).
 DR EMBL; AJ311904; CAC84565.1; --
 DR Genbank; HGNC:17111; ADAMS19
 DR InterPro; IPR002870; Rep_M12B_propep.
 DR InterPro; IPR001590; RepPolysin.
 DR InterPro; IPR000884; Tsf1
 DR InterPro; IPR000884; Tsf1 C2H2.
 DR InterPro; IPR004026; Zn_FpRptdse.
 DR Pfam; PF01564; Rep_M12B_propep. 1.
 DR Pfam; PF01421; RepPolysin; 1.

DR Pfam; PF00090; Tsp 1; 5.
 DR SMART; SM00209; TSE1; 5.
 DR PROSITE; PS02145; ADAM_MPRO; 1.
 DR PROSITE; PS00032; TSP1_5.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 SQ SEQUENCE 1207 AA; 134061 MW; AF36F6BF586FDE2 CRC64;
 Query Match 19.5%; Score 1018.5; DB 4; Length 1207;
 Best Local Similarity 31.7%; Pred. No. 2.1e-81;
 Matches 270; Conservative 110; Mismatches 296; Indels 175; Gaps 32;
 QY 167 LRLGPRHDFHMDRTSSSLVAPQFIVQTL-----GKTGTEKV--QTLPPEDFCFQCSL 219
 DB 167 LRLGPRHDFHMDRTSSSLVAPQFIVQTL-----GKTGTEKV--QTLPPEDFCFQCSL 219
 QY 220 ASHNSPSSH-----GGK-----PCGSGTRTLKLNQSKCPDROSVDFRAAQCAEH 263
 DB 220 ASHNSPSSH-----GGK-----PCGSGTRTLKLNQSKCPDROSVDFRAAQCAEH 263
 QY 264 NSRPSGRHRYKMKYQTVQVADLCKLYCIAEGDFFFSLSNKYKDGTPCSEDSRVNCDIGI 323
 DB 264 NSRPSGRHRYKMKYQTVQVADLCKLYCIAEGDFFFSLSNKYKDGTPCSEDSRVNCDIGI 323
 QY 269 PHRYTQKSEMEKVEKSA-LHSHYC--GI-----ISDK--GRP--RSKIAESG- 312
 DB 269 PHRYTQKSEMEKVEKSA-LHSHYC--GI-----ISDK--GRP--RSKIAESG- 312
 QY 324 CELSVSTSAHMPQFKEDLPILPDEYKSLRHRSLRSHRNEELNVELTVVVDKMMQ 383
 DB 324 CELSVSTSAHMPQFKEDLPILPDEYKSLRHRSLRSHRNEELNVELTVVVDKMMQ 383
 QY 313 -----RGRYVSKLPQSY-----NIETVVADPAMVS 339
 DB 313 -----RGRYVSKLPQSY-----NIETVVADPAMVS 339
 QY 384 NQGHENITVYTLNMSALFKDGLM----- 410
 DB 384 NQGHENITVYTLNMSALFKDGLM----- 410
 QY 340 YHGDAARFLLTLNMFNFQKSLGVQVNLKVLILLHETPPPELYGHGCKMLES 399
 DB 340 YHGDAARFLLTLNMFNFQKSLGVQVNLKVLILLHETPPPELYGHGCKMLES 399
 QY 411 -----QKGTGHDHAILLTGLDLSKWNKPCDGLGAPISQNSCKYBSCTINEDYGLGLAF 447
 DB 411 -----QKGTGHDHAILLTGLDLSKWNKPCDGLGAPISQNSCKYBSCTINEDYGLGLAF 447
 QY 400 FCKQHEERCKNDIHLENSINWEDMSVDAILLTRKDFCVHKDFCDIVGLAYLSM 459
 DB 400 FCKQHEERCKNDIHLENSINWEDMSVDAILLTRKDFCVHKDFCDIVGLAYLSM 459
 QY 448 CSKYBCTINEDTGLGATTAHSGHNFHMGHDEGCMCKSEG--NIMS--PTLAGRN--G 504
 DB 448 CSKYBCTINEDTGLGATTAHSGHNFHMGHDEGCMCKSEG--NIMS--PTLAGRN--G 504
 QY 460 CSERKCLIAEDNGLNLAFTIAHMGHNGINHDHPSG--ADGLHMSGEMIKQNLG 517
 DB 460 CSERKCLIAEDNGLNLAFTIAHMGHNGINHDHPSG--ADGLHMSGEMIKQNLG 517
 QY 505 VPSHSCSROVLYHKEFTQAQICLAD--OPKPVKEYKPEKLPGLDYDANTCKWQFGSKA 563
 DB 505 VPSHSCSROVLYHKEFTQAQICLAD--OPKPVKEYKPEKLPGLDYDANTCKWQFGSKA 563
 QY 518 DVSNSRCSKEDLEERLRSKASNCILLQTNPQSVNSVWVSKLPGLGWTITADECCQLLGLPLA 577
 DB 518 DVSNSRCSKEDLEERLRSKASNCILLQTNPQSVNSVWVSKLPGLGWTITADECCQLLGLPLA 577
 QY 564 KLCWLPKDKICKALCHRIGRK--CETKEMPAABGTICGHDMCRGQCVKYGDECPKPT 622
 DB 564 KLCWLPKDKICKALCHRIGRK--CETKEMPAABGTICGHDMCRGQCVKYGDECPKPT 622
 QY 578 SFQC--ENQHVICTGLCKVEGEKERTKLDPPMDGTDCLGKCKAGECTS--RTSAPEHL 635
 DB 578 SFQC--ENQHVICTGLCKVEGEKERTKLDPPMDGTDCLGKCKAGECTS--RTSAPEHL 635
 QY 623 HGHMSWSSWSPSCSRCTCGGYSHRSRLCTNPNKSHGKFCGSGTRTLKLNQSKCPDROSV 682
 DB 623 HGHMSWSSWSPSCSRCTCGGYSHRSRLCTNPNKSHGKFCGSGTRTLKLNQSKCPDROSV 682
 QY 636 AG--EWSLWSPSCSRCTAGISSRERK--FGLDSARDQNGPRKQYRCEINPPCAGLP 690
 DB 636 AG--EWSLWSPSCSRCTAGISSRERK--FGLDSARDQNGPRKQYRCEINPPCAGLP 690
 QY 683 DFRACCAHNSRRFRGRH--YKMKPYTOVEDQDLKLYCIAEGDFFFSLSNKYKDGTPC 741
 DB 683 DFRACCAHNSRRFRGRH--YKMKPYTOVEDQDLKLYCIAEGDFFFSLSNKYKDGTPC 741
 QY 691 GFRDMQCAQYSVRTSPFKILQWQ--AVLDSEKPCALFCSPVGEQFILLSEKVDGTS 748
 DB 691 GFRDMQCAQYSVRTSPFKILQWQ--AVLDSEKPCALFCSPVGEQFILLSEKVDGTS 748
 QY 742 SEDSRNVCDIGICERVCGDNVLGSDAIVEDVCGVGNNSACTIHRGLITKHHHTNYH 801
 DB 742 SEDSRNVCDIGICERVCGDNVLGSDAIVEDVCGVGNNSACTIHRGLITKHHHTNYH 801
 QY 749 GYQGLDITCANRCQKCGCDGLGLSLAREDHGCVNGKSGKCIKIGDF--NTRGAPYVEV 807
 DB 749 GYQGLDITCANRCQKCGCDGLGLSLAREDHGCVNGKSGKCIKIGDF--NTRGAPYVEV 807
 QY 802 VTIPSARSIRIYEMNVSYSIVRNALRYLNGHWTVDMPQRYKFSGTTFDYRRESYNE 861
 DB 802 VTIPSARSIRIYEMNVSYSIVRNALRYLNGHWTVDMPQRYKFSGTTFDYRRESYNE 861
 QY 808 LVIIPAGARRIKVVEEKPAHSYALRDA--GKQINSIDWKIEHSGAFNLAGTIVYVR--GL 865
 DB 808 LVIIPAGARRIKVVEEKPAHSYALRDA--GKQINSIDWKIEHSGAFNLAGTIVYVR--GL 865
 QY 862 PENLIATGPTNETL--IVELLFQGNPFGVAWEYS-----MPSLQTEKQPPAQPSYVAIVR 915
 DB 862 PENLIATGPTNETL--IVELLFQGNPFGVAWEYS-----MPSLQTEKQPPAQPSYVAIVR 915
 QY 866 MEKISAKGPTAPLHLLVLLFQDQNYGLHYEYITPSDFLPENOSSKAP--EPLFWHTHTS 923
 DB 866 MEKISAKGPTAPLHLLVLLFQDQNYGLHYEYITPSDFLPENOSSKAP--EPLFWHTHTS 923
 QY 916 --SECSVSCGG 925
 DB 916 --SECSVSCGG 925
 QY 924 MEDCDATCGG 934
 DB 924 MEDCDATCGG 934

```

RESULT 10
Q19791
ID Q19791 PRELIMINARY; PRT; 967 AA.
AC Q19791;
DT 01-OCT-2002 (TRENBLREL. 22, Created)
DT 01-OCT-2002 (TRENBLREL. 22, Last sequence update)
DT 01-NAR-2003 (TRENBLREL. 23, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Strausberg R.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC036515.1;
DR InterPro; IPR006586; ADAM_Cysteine.
DR InterPro; IPR001128; Cytochrome_P450.
DR InterPro; IPR002870; Pep_M12B_propep.
DR InterPro; IPR001590; Reprolysin.
DR InterPro; IPR000884; TSPI.
DR InterPro; IPR006025; Zn_MTPeptdase.
DR Pfam; PF01562; Rep_M12B_propep. 1.
DR Pfam; PF01421; Reprolysin. 1.
DR Pfam; PF00090; tsep_1; 2.
DR SMART; SM00608; ACR; 1.
DR SMART; SM00209; TSPI; 3.
DR PROSITE; PS00215; ADAM_MEPRO; 1.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
DR PROSITE; PS00092; TSPI; 3.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Hypothetical protein.
SQ SEQUENCE 967 AA; 105387 MW; FFD399674201C3D CRC64;

Query Watch 15.8k; Score 984.5; DB 4; Length 967;
Best Local Similarity 30.9%; Pred No. 1.6e-78;
Matches 270; Conservative 104; Mismatches 326; Indels 173; Gaps 35;

QY 157 VAVSEV-----SLHLKGRPHDFHDLRTSSLSVAFPIVTLG-KGTGKSVQTLPP 209
DB 160 LVVPELAPGHOTRLHLAFQQLDLRLPDSGFLAGFTLQNVGRKSGSET--PLPE 117
QY 210 EDF--CFYQGLSHRNPSHGKFCGEGSTRTLUKLSN-----OKCPDSDVDPRAAQCAH 263
DB 118 TDLAHCFTSGTVNGDPSSAA--ALSGLGRCGAFYLLGEAYFIQPLPAASERLATAAPE- 175
QY 264 NSRRFRGRHYKWFYVQVEADLCKLYCAEGFOFFFSLSNKKVDGTPCSDSRNVCIDGI 323
DB 176 -----KPPAPLQ-----PHLLRNQSGD-----VGGT 197
QY 324 CELSVVSTSAH-----MPQPPEDLFLIPDKVSKLRRKSL 360
DB 198 C--GVVVDPEPTGKAETEDSDGTEGDEGPQWSPDPAQGVQGTGTGS-IRKR-F 253
QY 361 LRSRHEENLVTLVVVKKMKNQNGHENITVTVTLNVMVSALF----- 406
DB 254 VSSR-----YVETMLVADQSMABFHG-SGLKHYLLTFSVAARLYNKPISRNVSUVVVK 308
QY 407 -----DGLMG-----KQGRHDHAILTLGLDTCGWRNEP 435
DB 309 ILVNHDEKQGEPTVSTNAALTIRNFWKQKQHNPFSDRAEHYDTALFTQDLG--SQT 366
QY 436 CDTLGAPISOMGKSVRSTTNDTGLGAPTAHSGHNFMIHDSGNMCKSEG--- 492
DB 367 CDTLGADVGVTCDSRSCSVIEDDGLQAFTTAHELGHVFNPHD-DAKQASLNGVQ 425
QY 493 --NIMSPTLAGRNGVSPSCSYLHKFLSTAAQICLADQPKPKYKPEKLGELYD 550
DB 426 DSHWASMLSNLHOSQWSPSCSYAMITSLFNGHGLCLMD--XPQNP:CLPGLPFTSYD 483
QY 551 ANTQCKWQFGKAKLCLMDLFFKKDICKALWCHIRGRK-----CEKFNPAEAGTICGHDMW 605

```

RESULT 11

```

Q19791
ID Q19791 PRELIMINARY; PRT; 2165 AA.
AC Q19791; Q27524 (TRENBLREL. 01, Created)
DT 01-NOV-1996 (TRENBLREL. 08, Last sequence update)
DT 01-NAR-2003 (TRENBLREL. 23, Last annotation update)
DE P55H8.3 protein.
GN P55H8.3
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OC NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Gajadasy S.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M., Bonfield J., Burton J., Connell M., Cooper T., Coulson A., Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L., Jones M., Kershaw J., Kirschen J., Laister N., Latreille P., Lathigra J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M., Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R., Smalton N., Smith A., Sonhammer E., Staden R., Sulston J., Thierry-Mieg J., Thomas K., Vaughan M., Vaughan K., Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J., Woldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans."
RT Nature 368:32-38(1994).
DR EMBL; Z69361; CAA93288.1;
DR EMBL; Z69360; CAA93287.1; JOINED.
DR EMBL; Z69361; CAA93287.1; JOINED.
DR MEROPS; M12.135; -.
DR WormPeP; F25H8.3; CE05729.
DR InterPro; IPR001590; Reprolysin.
DR InterPro; IPR000884; TSPI.
DR InterPro; IPR006025; Zn_MTPeptdase.
DR Pfam; PF01421; Reprolysin; 1.
DR Pfam; PF00090; tsep_1; 14.

```


SMART, SMO209, TSPL, 18
 DR PROSITE, PS00215, ADAM, MEPRO, 1.
 DR PROSITE, PS0092, TSPL, 15
 DR PROSITE, PS00142, ZINC, PROTEASE, 1.
 SQ SEQUENCE 2165 AA; 24497 MW; FCC3D8A8AA9C4888 CRC64;

Query Match 18.4%; Score 965.5; DB 5; Length 2165;
 Best Local Similarity 27.9%; Pred. No. 2.7e-76;
 Matches 264; Conservative 134; Mismatches 333; Indels 215; Gaps 31;

QY 122 LHVTQYDYLVAEVDHGRDVGSHLHMHQRRRAVASE--VESLHRLKRGPRDHMO 179
 DB 66 IHVIDSHHIVRDSYGRG--KEDVTSTDRRLRGQVARDGCHACHRLRSLSDAVYIVH 122
 QY 180 LRT-----SSSLVAFGFIQTLGKTGTSVQTLPPDEFCFYQSLRS--HR 223
 DB 123 LHRNQIIPDSHNSVPHPSNSNAPMVLVLDSEEVGRGMSRTDPD--CIYRAHVKGQV 180
 QY 224 NSPSHGKFCGSGTFLKLCNSQK-----CPROSDVDFRAAQ 260
 DB 181 HS-----IVNLICDSEGLYGLMALPSGHIVPEIISNGTEHOG--ASRH 223
 QY 261 AEHNSRRFRGRHYKPYTOVEADCLCYIAEGDFFSLSNKVKDGTGCSBDSRNCV 320
 DB 224 BQHLVVRKQDVRHEK-----SPDHLNSTSVNETTT----- 253
 QY 321 DGICELSVVSTSAHMPQPKEDLPILPDYKSLRHKRSLRLSRHNEELNVELTVVDK 380
 DB 254 -----VAT-----WQDMEDVIERKARSRAANSWQHIVVELVWADTK 291
 QY 381 MNQNHGHEMITTYVILNMVSAFLKDGUMGKO-----GTR----- 416
 DB 292 MEYHG--RGEDVYLFTVAYRHSGLASINVVVVKLVLTKEWAPRITQNAQOT 350
 QY 417 -----HDAHLITGLDLSCHWKEPCDILGAPISGKYSRCT 455
 DB 351 LQDFCRWQVYNDPDSVQYHEDVALUTKDCRSQK--CQNLGLASLGTMDCKSCA 409
 QY 456 INEPTGLGLATIAHSGHNGFMHGE-----GNMCK-----KSEGIN 495
 DB 410 IIEDNGLSAAFTIAHELGHVFSIFDHDERKCSYVMPVKVCKFQSTKDKTQFQNNFHM 469
 QY 496 SPTLAGRNVFSWSPSCROYLHKFL--STAQAICLADQPKPKVEYK--YPEKLPGELYDA 551
 DB 470 APTLEYNTHPSWSPSCSAGMLERFLNRRGQTCLDFQFVERRYEDVFRDEPKKYDA 529
 QY 552 NTQCKWQFGEKAKLMDPKKDKICAKLMC-----HRIGKCTKFPAAEGTIC--GHDM 604
 DB 530 HQOCKVFVFPASLCTPY--NPTCRRLCATFYGSQMG--CRQTHFWADGTFCDESRSM 584
 QY 605 WCRGGCVKRYGDEGKPTGHWSWSPSCSRCTCGGVSHRSRLCTNPKPSHGKFKFCEG 664
 DB 585 FCHHGACVRLAPESLTKIDGQWDMWSGECSTRTCGGVQKGLRDCSPKPRNGGKYCVG 644
 QY 665 SRTLKLCNSQKSPDSVFRFAAQCAHNSRRF-----RGRHYKNAP--YTOVEDDQDLKL 718
 DB 645 QREYRSCNTQECWDTQPYREVQCSEFNKNDIGIQVASTNTHWYKYANVAPNERCKL 704
 QY 719 YCTAAGDFFPSLSNKVKDGTGCSBDSRNCVIGICERVCGCNVLGSDAIEDVCGVCGN 778
 DB 705 YCRUSSGAFLYLRDKVDVDPGTCRNGDDICVAGACWAGCDHLSTLRDKCGVCGGD 764
 QY 779 NSACTIHRGLTKHHHTNYHVMVTIPSGARSIRIYE-----MNSTSYISVRNALRRY 833
 DB 765 JSSCKVVVKGTNE--OGTFGEVWKIPAGSANIDIRQKYNKMKEDDNYLSRAANGSFL 823
 QY 834 LNHWTVDMP--GRYKSGTTFDYRSNEPENLIATGPTNETLIVELLFQGRN--PGVAWE 891
 DB 824 LNHGFOVSLARCAQTAQTVLIEGSDAILERNGTPIRSDIYVHVLSVSGSHPPDISYE 883
 QY 892 ISMPLU--GTEKQPAQPSYTWAIVR--SECSVSCGGRCLFILLE 934
 DB 884 YWYANPNAVNPFIPISSALLVLRVTDVTELDACRCQSQSQKMLCD 929

RESULT 12

QY 371 VETLVVDVKKMQRHGHENITTYVILNMVSAFLKDGUMGKO-----GTR----- 410
 DB 52 VETLVVDVKKMQRHGHENITTYVILNMVSAFLKDGUMGKO-----GTR----- 410
 QY 411 -----CKGTHDHAILLTGLDLSCHWKEPCDILGAPISGKYSRCT 445
 DB 111 EVTSVNAITLRFQWQKQNPDPDRDEHYDTALTFRDQCG--AQTCDLGLVADVG 168
 QY 445 GHSKYSRCTINEDTGLATIAHSGHNGFMHGE-----GNMCK-----KSEGIN 495
 DB 169 TICDPSESCSVIEDDGLQAAFTTAHELGHVENPHD--QAKQCSASINGVRDSDHMA 227
 QY 501 GNGVFSWSPSCROYLHKFL--STAQAICLADQPKPKVEYK--YPEKLPGELYDA 551
 DB 228 NLDSPQSPSCSAYMITSLDNGHGECLMD--KPSPIQLPSDLPTGLYDANRQCF 285
 QY 561 EKAKLMDLDFKKDKICAKLMC-----HRIGKCTKFPAAEGTIC--GHDM 604
 DB 286 EESKHC--PDAASTCTTLMC--TGTSGLLVCQTKHFWADGTGSGEGRMCVNGK 341
 QY 616 DEG--PKPTHGSDMSWSPSCRTCGGVSHRSRLCTNPKPSHGKFKFCEG 664
 DB 342 DRKHFTDPVHSGMGPMPGDCSTCGGVVYTNRECLNPPVNGKCYCEGKAVR 401
 QY 674 SQKPC--RDSVFRFAAQCAHNSRRF-----RGRHYKNAP--YTOVEDDQDLKL 718
 DB 402 IEDCPNNNGKTFREEQCAHNEFSKASFGSPAVENWPKYAGVSPKDRCKL 704
 QY 728 FFSLSNKVKDGTGCSBDSRNCVIGICERVCGCNVLGSDAIEDVCGVCGN 778
 DB 462 FVLVQKPVVDGTGCSBDSRNCVIGICERVCGCNVLGSDAIEDVCGVCGN 778
 QY 788 LYTHHTNYHVMVTIPSGARSIRIYE-----MNSTSYISVRNALRRY 833
 DB 522 SVTS--ARPGVHDIPTGTATVIEVKORNGSRNNGSFLAKAADDTVLNGDFT 579
 QY 843 PCR--YKFGSTTFDYRSNEPENLIATGPTNETLIVELLFQGRN--PGVAWE 891
 DB 590 LEQDITYKSGVLSVSGSAALEHRSFSLPELTIQVLTGVNLRPARKITYFVK--K 637
 QY 900 EKQPAQPSYTWAIVR--SECSVSCGGRCLFILLE 934
 DB 638 KESFNAIPTFSWVIEWEGECSKSCQGG 665

RESULT 13
 Q8K384 PRELIMINARY; PRT; 497 AA.
 AC Q8K384
 DT 01-OCT-2002 (TEMBLrel. 23, Created)
 DT 01-MAR-2003 (TEMBLrel. 23, Last sequence update)
 DE Similar to a disintegrin-like and metalloprotease (reprolysin type)
 DE with thrombospondin type 1 motif, 4.
 OS Mus musculus (Mouse)
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 RL EMBL; BC027773; AAH27773.1; -
 DR InterPro; IPR006586; ADAM cysteine.
 DR InterPro; IPR001590; Reprolysin.
 DR InterPro; IPR000884; TSP1.
 DR InterPro; IPR006025; Zn.MTpeptidase.
 DR Pfam; PF01421; Reprolysin; 1.
 DR Pfam; PF00090; tsp.1; 1.
 DR SMART; SM00608; ACR; 1.
 DR SMART; SM00209; TSP1; 1.
 DR PROSITE; PS02015; ADAM_MEPRO; 1.
 DR PROSITE; PS00992; TSP1; 1.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 KW Integrin; Protease; Metalloprotease.
 SQ SEQUENCE 833 AA; 90097 MW; 4CECB3DFBC3AA619 CRC64;
 Query Match 16.7%; Score 872.5; DB 11; Length 833;
 Best Local Similarity 34.4%; Pred. No. 1.3e-68;
 Matches 216; Conservative 77; Mismatches 330; Indels 105; Gaps 23;
 QY 371 VETLVVDDKMMQNGHGHENITVLTILNNVSAALFXD----- 407
 DB 216 VETLVVADDDKMAFHG-TGLKRYLLVVAAKAKAFHPSIRPNLVVTRVLVILGSSQGS 274
 QY 478 -----CLMGKDGTRHDH---AILLTGLDICSWNKPEPDTLGPAPIS 445
 DB 275 PQVGPSSAQTLASCTWQRLNTNDSDFEDTALIFTRQLDGVST--CUTLGMADVG 332
 QY 446 GNCYSKYSTINEDTGLGLAFTIAHSGNPFQTHD-----GEINMCKSEGNHMS 496
 DB 333 TVCDPAPRACVAVDQSGAFNAHELGHVFNHLDNSKFNCTNUNGQGG---SSRHVPA 388
 QY 497 PTLAQRNVGPSFCSQRYLHKFLSTAQAILAQAPQPKVKEVPEKLDELVDANTOCK 556
 DB 389 PVAHVDPPEPAPSCSARITDPLDNGYGHCLLO--KPEAPLHLPATFFKQYDNDQQQ 446
 QY 557 MQCFEKAICMLDFKDIKCALVK--HRIGR-KCETFKFPRAEGTIGCHMCKGCGCYK 613
 DB 447 LTFGPDSSHC--POLPPPCALKSCSHLNGHAMQTKVSNADGTPCGSSQACVGGRCILH 504
 QY 614 YCD-EGKPTGHGWSMDSPSCPCRTCCGGVYSHRSFLCTNPKPSHGKFCESSTRITLKL 671
 DB 505 VDQLKDFNVFQAGGQGWPGWGDSCRTCCGGVQVFSRDCTRPVPRNGSGVCEGRRTRFRS 564
 QY 672 CNSQKCFRDS-VDFRAQCAEHNSR-----RFRGRHYKWK-PYTOVEDODLCKLYCIABG 724
 DB 565 CNTENCCHGSALTFRREQCAATNHRITDLFKSPFG-PMDVPRYRTGVAPRQDCKLTQCARA 623
 QY 725 PDPFFSLSKNVKDGTPCSDSDSNVCDIGICERVQCNVLGSDVADEVQVCNGNNSACTI 784
 DB 624 LGYVYVLEPRVADGTPCSDTSVVCVQGRCTHAGCDRIIGSKKKFKDKCMVCGDGRCSK 683
 QY 785 HRGLYTKHHHTNYHMTVTPSGARSIRIVEM--NVSTSYISVRNALRYVLNGHWT-- 839
 DB 684 QSGSPKKFRY--GYSDVVTIPAGATHILVRQGGSGLKIYLAJLKLSGDSVALNGEYTL 741
 QY 840 -----VDWFG--RYKFSGTTTFDYRSYNPEPNIATGPTNETLIVELLFQG--RNPQVAM 890
 DB 742 PSPTDVLVFGAVSLRYSAGAT-----AASETFLSGHGPLADPLTLOVLVAGNPQARLY 794

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2003, 23:39:29 ; Search time 53 Seconds
759.999 Million cell updates/sec

Title: US-09-981-151A-8
Perfect score: 5236
Sequence: 1 MKPRARGMGLAALMLLAQ.....LENACQSPATYALAFLES 952

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues
Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum March 0%
Maximum March 100%
Listing first 45 summaries

Database : Issued Patents AA*
1: /cgm2_6/prodata/2/iaa/5A.COMB.pdp:*
2: /cgm2_6/prodata/2/iaa/5B.COMB.pdp:*
3: /cgm2_6/prodata/2/iaa/6A.COMB.pdp:*
4: /cgm2_6/prodata/2/iaa/6B.COMB.pdp:*
5: /cgm2_6/prodata/2/iaa/PCTUS.COMB.pdp:*
6: /cgm2_6/prodata/2/iaa/backfiles1.pdp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3846.5	73.5	1224	4	US-09-930-872-4
2	1543.5	29.5	491	4	US-09-930-872-2
3	1198	22.7	859	4	US-09-369-364A-5
4	1179	22.5	997	4	US-09-369-364A-7
5	1057.5	20.2	1081	4	US-09-369-364A-17
6	1056	20.2	1882	4	US-09-369-364A-13
7	991	18.9	874	4	US-09-369-364A-15
8	984.5	18.8	967	4	US-09-130-491-2
9	958.5	18.3	727	4	US-09-445-023A-12
10	934	17.8	727	4	US-09-445-023A-1
11	933	17.8	1205	4	US-09-491-522-11
12	923	17.5	1211	4	US-09-491-522-5
13	903	17.2	608	4	US-09-130-491-13
14	895.5	16.9	930	4	US-09-369-364A-2
15	878	16.8	930	4	US-09-122-126B-15
16	878	16.4	837	4	US-09-634-286A-15
17	856.5	15.4	837	4	US-09-122-126B-2
18	856.5	15.4	837	4	US-09-634-286A-2
19	823.5	15.7	905	4	US-09-369-364A-9
20	778.5	12.9	551	4	US-09-130-491-16
21	675.5	12.9	518	4	US-09-369-364A-22
22	670	12.6	481	4	US-09-130-491-8
23	407	10.0	566	4	US-09-491-522-7
24	451.8	8.6	525	4	US-09-369-364A-21
25	323	6.2	450	4	US-09-369-364A-19
26	270.5	5.2	416	4	US-09-608-790-1
27	268.5	5.1	751	2	US-08-836-443-3

28	245.5	4.7	802	4	US-09-632-098-2
29	245.5	4.7	812	4	US-09-632-098-4
30	244.5	4.7	814	3	US-09-813-819-4
31	244.5	4.7	814	4	US-09-920-048-4
32	232.5	4.4	855	3	US-09-813-819-2
33	232.5	4.4	855	4	US-09-920-048-2
34	232	4.4	1172	1	US-08-313-288B-19
35	228	4.4	1170	1	US-08-313-288B-20
36	217	4.1	748	1	US-08-920-234-2
37	215	4.1	748	3	US-09-030-335-9
38	215	4.1	769	1	US-08-243-542-4
39	215	4.1	769	1	US-08-477-407-4
40	215	4.1	769	1	US-08-484-355-4
41	212.5	4.1	529	2	US-08-836-442-3
42	208	4.0	245	4	US-09-369-364A-11
43	205	3.9	441	3	US-08-985-526-3
44	204.5	3.9	464	3	US-09-411-329C-14
45	203	3.9	670	1	US-08-243-542-3

ALIGNMENTS

RESULT 1
US-09-930-872-4
; Sequence 4, Application US/09930872
; Patent No. 6448388
; GENERAL INFORMATION:
; APPLICANT: Hilbun, Erin
; TITLE OF INVENTION: No. 6448388 Human Proteases and Polynucleotides Encoding the
; FILE REFERENCE: LEX-0219-USA
; CURRENT APPLICATION NUMBER: US/09930872
; PRIOR FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: US 60/225,852
; PRIOR FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1224
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-930-872-4

Query Match	73.5%	Score	3846.5	DB	4	Length	1224
Best Local Similarity	75.4%	Pred. No. 0					
Matches	738	Conservative	25	Mismatches	73	Indels	143
Gaps	13						
QY	1	MKPRARGMGLAALMLLAQVAEQVFGSRHQGNRGSGQLEASPPRLSRFLRTAMS	60				
DB	1	MKPRARGMGLAALMLLAQVAEQVFGSRHQGNRGSGQLEASPPRLSRFLRTAMS	60				
QY	61	ELFSAGTCVVRHGTSGSAMPERPASSSTRGAAGLDGKGRDMBAGNHRSGQNTGTENG	120				
DB	32	FAAAA-----PCSPSPRPPPPAEPG-----MNEKG-----	58				
QY	121	TLHYLTQDLSAYVDHEDGYVSHMEHQRRRAVAVSEVSHLRKGRPHDFHML	180				
DB	59	-----EYDLVSAYVDHEDGYVSHMEHQRRRAVAVSEVSHLRKGRPHDFHML	112				
QY	181	RTSSSLVAFGRIVTGLTKGTNSVQTLPPDFCYQGSLSHRNSPSHGKFCBEGSTRTL	240				
DB	113	RTSSSLVAFGRIVTGLTKGTNSVQTLPPDFCYQGSLSHRNS-SVALSTCQGLSGMI	171				
QY	241	KLCNOKCPROSDVFRAAQCAEHNRRFRGR-----HYKWKPYTQVEA-DLCKLYCI	291				
DB	172	R-----TEADYFLRFLPSHLSWL-GRAAQSSPSHVLYKSTEPHAGSEVLVT	222				
QY	292	AGEGDFEFS--LSNKKXGDTGCCSDSRNVCIDIGICELSVVSTASHMPQPPKEDFLPDE	349				
DB	223	SRTVELAHQPLHSSDLRLGLP---QKHFC-----GRKKYNPPQPPKEDFLPDE	270				
QY	350	YKSLRHKRSLLSRHNEELNVETLVVVDKMMQNHENITTVYVLTILNMVSAFKD--	407				

```
Db 271 YKSLRHKSLRSHRNEELVETLVVVKQKQKHGHEHNTYVTLILNMVSAFLDGT 330
|||
Qy 408 -----GLMGKDGTRDHAILLTGLD 427
|||
Db 331 IGGNINIAIVGLLEDEQGLVISHHAADHTLSSFCQWQSGLMGKDGTRDHAILLTGLD 390
|||
Qy 428 ICSWNEPCDTLGFAPISGCMKSKYRSCITNEDTGLGLAFTIAHESGHNFGMHDGEGNM 487
|||
Db 391 ICSWNEPCDTLGFAPISGCMKSKYRSCITNEDTGLGLAFTIAHESGHNFGMHDGEGNM 450
|||
Qy 488 KXSEGNIMPTIAGRNQVFNQSPCRQYLHKFLSTAQAICLADQPKPVXYKYVPEKLPGE 547
|||
Db 451 KXSEGNIMPTIAGRNQVFNQSPCRQYLHKFLSTAQAICLADQPKPVXYKYVPEKLPGE 510
|||
Qy 548 LYDANTQKQWQGEKAKLWMLDKXDKICAKLMCHIRGRKCTKFWPAAEGTICGDMWCR 607
|||
Db 511 LYDANTQKQWQGEKAKLWMLDKXDKICAKLMCHIRGRKCTKFWPAAEGTICGDMWCR 570
|||
Qy 608 GQCVKYGDEGPRFTHGHSNDSWSPCSRTCGGVSHRSLCTNPKPSHGKFCGEGSTR 667
|||
Db 571 GQCVKYGDEGPRFTHGHSNDSWSPCSRTCGGVSHRSLCTNPKPSHGKFCGEGSTR 630
|||
Qy 668 TLKLSNKKVKGDTGPPCSGDSRNVCI DGI CERVGDNVLSGDAVEDVCGVGNNSACTIHRG 727
|||
Db 631 TLKLSNKKVKGDTGPPCSGDSRNVCI DGI CERVGDNVLSGDAVEDVCGVGNNSACTIHRG 690
|||
Qy 728 FFLSNKKVKGDTGPPCSGDSRNVCI DGI CERVGDNVLSGDAVEDVCGVGNNSACTIHRG 787
|||
Db 691 FFLSNKKVKGDTGPPCSGDSRNVCI DGI CERVGDNVLSGDAVEDVCGVGNNSACTIHRG 750
|||
Qy 788 LTKKHHHTNOYHMTVTPSGARSIRIYEMNVSTSVIYRNALRYILNGHWTDWPGRYK 847
|||
Db 751 LTKKHHHTNOYHMTVTPSGARSIRIYEMNVSTSVIYRNALRYILNGHWTDWPGRYK 810
|||
Qy 848 FSGTTFDVRKSYNEBENJATGCTNETHVLELLFCQNGCVAMEYMPRLGTXKQBPAP 907
|||
Db 811 FSGTTFDVRKSYNEBENJATGCTNETHVLELLFCQNGCVAMEYMPRLGTXKQBPAP 870
|||
Qy 908 SYTVAIVRSCVSCGGR 926
|||
Db 871 SYTVAIVRSCVSCGGR 889
|||
```

RESULT 2

```
US-09-930-872-2
; Sequence 2, Application US/09930872
; Patent No. 6448388
; GENERAL INFORMATION:
; APPLICANT: Fridge, Carl Johan
; TITLE: INVENTION: 6448388el Human Proteases and Polynucleotides Encoding the Sa
; FILE REFERENCE: EX-0210-USA
; CURRENT FILING DATE: 2001-08-14
; PRIOR FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO: 2
; LENGTH: 491
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-930-872-2
```

```
Query Match 29.5%; Score 1543.5; DB 4; Length 491;
Best Local Similarity 57.7%; Patent No. 18-129;
Matches 333; Conservative 26; Mismatches 75; Indels 143; Gaps 13;
Qy 1 MKPRARGHGLAALWMLAQAQVSPQRSHQSGQLEAPRLLSGPRRLTAS 60
|||
Db 1 MKPRARGHGLAALWMLAQAQVSPQRSHQSGQLEAPRLLSGPRRLTAS 60
|||
```

```
Qy 61 PLFSAGTCVVRHGRSGANEPERPASSSTRGAAGLCKGRDMDEAGNHRSGQQTNTGTENQ 120
|||
Db 32 PAAAA-----PSPSPVPRPPPAERPG-----WMKSG-----56
|||
Qy 121 TLHVLTOYDLVSAVEVDHGDYVSHETMHQRRRAVAVSEVESLHLRLKGRPHDFMOL 180
|||
Db 59 -----EYDLVSAVEVDHGDYVSHETMHQRRRAVAVSEVESLHLRLKGRPHDFMOL 112
|||
Qy 181 RTSSSLVAGFGIVOTLGTGTGKSVOTLPPDECFYQGSLSHRNSPSHGKFCGEGSTR 240
|||
Db 113 RTSSSLVAGFGIVOTLGTGTGKSVOTLPPDECFYQGSLSHRNS-SVALSTSCQGLGMI 171
|||
Qy 241 KLCNSQKCPDSVDVFRAAQCAHNSRRFRGR-----HYKWPYQVEA-DLCKLYCI 291
|||
Db 172 R-----TEADYFLRLPSPHLSWKL-GRAAQSPSHVLYKRSSTEPHAFGASEVLVT 222
|||
Qy 292 AEGDFDFFS--LSNKKVKGDTGPPCSGDSRNVCI DGI CELSVVSTVSAHMQPKEDLPILOE 349
|||
Db 223 SRTMELAHQPLHSSDLRLGLP-----QKHFC-----GRKKYMPQPPKGLFIPDE 270
|||
Qy 350 YKSLRHKSLRSHRNEELVETLVVVKQKQKHGHEHNTYVTLILNMVSAFLDGT 407
|||
Db 271 YKSLRHKSLRSHRNEELVETLVVVKQKQKHGHEHNTYVTLILNMVSAFLDGT 330
|||
Qy 408 -----GLMGKDGTRDHAILLTGLD 427
|||
Db 331 IGGNINIAIVGLLEDEQGLVISHHAADHTLSSFCQWQSGLMGKDGTRDHAILLTGLD 390
|||
Qy 428 ICSWNEPCDTLGFAPISGCMKSKYRSCITNEDTGLGLAFTIAHESGHNFGMHDGEGNM 487
|||
Db 391 ICSWNEPCDTLGFAPISGCMKSKYRSCITNEDTGLGLAFTIAHESGHNFGMHDGEGNM 450
|||
Qy 488 KXSEGNIMPTIAGRNQVFNQSPCRQYLHKFLSTAQA 524
|||
Db 451 KXSEGNIMPTIAGRNQVFNQSPCRQYLHKFLSTAQA 487
|||
```

RESULT 3

```
US-09-369-364A-5
; Sequence 5, Application US/09369364A
; Patent No. 6391610
; GENERAL INFORMATION:
; APPLICANT: Apce, Suneel
; APPLICANT: Hurskainen, Tiina L.
; APPLICANT: Hirohata, Satoshi
; TITLE: INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
; FILE REFERENCE: 26473/4607/18-30-00
; CURRENT FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO: 5
; LENGTH: 859
; TYPE: PRT
; ORGANISM: Homo sapiens ADAMTS-6
; NAME/KEY: MOD RES
; LOCATION: (453)
; OTHER INFORMATION: Xaa = L
US-09-369-364A-5
```

```
Query Match 22.7%; Score 1188; DB 4; Length 859;
Best Local Similarity 31.4%; Pred. No. 1.8e-97;
Matches 284; Conservative 132; Mismatches 289; Indels 200; Gaps 25;
```

```
Qy 108 HRSQQTWTGTENQTLHVLTOYDLVSAVEVDHGDYVSHETMH--HQRRAV----AVSE 161
|||
Db 22 HSDRLSYSGEEFLTYLHYQLPIRVDQNGAFLFTVVDKXHSRRSRKMDPQQA 81
|||
Qy 162 VESLHLKGRPHDFMOLATSSSLVAGFGIVOTLGTGTGKSVQTLPE-----DRCFY 215
|||
Db 82 VSKLFFKLSAYGKIFHLNLTNDFVSKIFTVYVKGOG-----PQWKHDFLNDCHY 133
|||
```

```

QY 216 QCLSRHNSPSHGKFCBEGSTRTLKLCNSQKCPDSVDFRAAOCAE-HNSRRRFRGRHYK 274
DB 134 TGYLQDOR-----STTKVALS-----CVGLAGVIATEDERYF 166
QY 275 WKPYQVQADLCKLYCIASGPPDFFSLSNVKADGTP-----CSEDSRNVICDIGELSV 328
DB 167 TEPLKNTED-----SKHFSY-----ENGHPHYVYKKSALOORHLVDHSHGVSVD 211
QY 329 VSTSAHWPPOPKEDFILDEYKSC---LRHRSLLRSRHEELNVETLVVVVDKQMHQ 385
DB 212 FTRSGKFWMLNDTFLFIHQVQINTHHRRQRSV---SIESFVETLVVADKQWGVYH 266
QY 386 GHEMITVYVLTALNMSALFKDGLMG-----TRHDHAILLGLDICSWNKPCDTIGFAPISGCKYRS 453
DB 267 GRKOIEHYLSVMVIAKLYDRSSGLNVNVIIVARLVILTEDQPNLEINHADKSLDSCF 326
QY 412 -----KDG-----TRHDHAILLGLDICSWNKPCDTIGFAPISGCKYRS 453
DB 327 KWOKSILSHQSDGMTIPENGIAHNDNAVILTYDICTYKNKPCGTGLASVAGNCPERS 386
QY 454 CTINEDYGLAFTIAHESGRNFOWHDEGNMCKYSEGNINMSTPLAGNNGVSNFSCSR 513
DB 387 CSINEDIJGSAFTIAHEIVHNFQWMDHGNSC-----GR-----422
QY 514 QYUHKFLSTAQAICLAQDPKPKYKPEKLPBELYDANTQCKWQGEKAKLML-----568
DB 423 -----KYNQK-----QNYGSHYCEYQ--SFFLVCLQSRXHH 452
QY 569 DFKXDKCALMCHRIGRKCTEYFPAABOTICQ-----HDMKRGQCYKQDEOPKETHG 624
DB 425 QUREVCEJMLCKSFKSRVCTYIPAREGTLCTGNIEMKWCYQDCVFFG-TWQSDIG 511
QY 624 HNSWMSWQSCPCGCGHSHSRCLCTNPKSHGKSCGSTRTLKNSQKCPDSVDF 684
DB 512 GNGFSLWQECSTICGGVSSSLHDCSPAPBEVETLGERKYSCHITDEPLGSRDF 571
QY 685 RAAQCAHNSRFRGRHYKPKYTOVEDODLCKLYCIASGPPDFFSLSNVKADGTPCSD 744
DB 572 REKQDFDPMFPKQYVYKPKYTCGGVKE-CALNCLAGEINFTYERAPAVIDGTQCAD 630
QY 745 SRNVICDIGERVCNDVLGSDAVEDYCGVGNHNSACTIHRGLYTKHHHTNQYKXVTI 804
DB 631 SLDICINSECKHVGCDNIGSDAREDCRVCVCGSGSTCDALIEGFNDLSLPGCYWEVQI 690
QY 805 PGARSIRIYENVNSTYSIVSNALRYVINGHTVDMPCRYKSCPTTFOVERSYMEREN 864
DB 691 PRGSVHIEVREAVKSNVYIALKSEGGDYINGANTIDWPKFQVAGTAFHKKRPTDEPS 750
QY 865 LIATGPNETLIVELLQGRNFGVAVYSNP--RLGTGKOPPAQPSYTAIVR-SECSVS 921
DB 751 LEALGPTSENLIVWVLLQCNLGIRYKFNVPITRTGSDN---EVGFTNHPQWSECSAT 807
QY 922 CGGGR 926
DB 808 CAGGR 812

```

RESULT 4

US-09-369-364A-7

Sequence 7, Application US/09369364A

Patent No. 6391610

GENERAL INFORMATION:

APPLICANT: Apte, Suneel

APPLICANT: Hurskainen, Tiina L.

APPLICANT: Hirohata, Satoshi

TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases

FILE REFERENCE: 26473/4007/10-30-00

CURRENT APPLICATION NUMBER: US/09/369,364A

NUMBER OF SEQ ID NOS: 31

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 7

LENGTH: 997

TYPE: PRT
ORGANISM: Homo sapiens ADAMTS-7
US-09-369-364A-7

Query Match 22.5%; Score 1179; DB 4; Length 997;
Best Local Similarity 30.3%; Pred. No. 1.5e-96;
Matches 292; Conservative 125; Mismatches 307; Indels 240; Gaps 26;

```

QY 129 DLVSAYEDVRHGDYVSHSEMHHRRAVAVSEVESLHLKPRHOFHMDLRTSSSLVA 188
DB 44 DIVHEFVVDAGGSFLSYELMPRALFKRDVSRRDAPAFYELQYGRRLRENFUTANQHLA 103
QY 189 PGFIVOTLKGTKSVQVTPDFECFYQGSLSHRNSPSHG---XFCGGSTRTLKLCN 244
DB 104 PGFVSETRRGGGLGNAHRAHTPACHLLGEV---ODPELGGGLAAIACDGLKGVQLSN 160
QY 245 SOKC--PRDSVFPRAOACAH-----NSRRFRGRHYKWKPYT---QVEADLCKLYCIAE 293
DB 161 EYFTEPLDSAPARQAQSHVYVYKQAPRLAQRGSSASPTCGVQVPEL-----212
QY 294 GFDFFFSLSNVKADGTPCSEDSRNVICDIGELSVSTSAHWPQPKEDFILDEYKSC 353
DB 213 -----ESRR-----222
QY 354 LRHKSLLRS-HR--NEELNVETLVVDDKQMHQHENITVYVLTALNMSALFKDGL 409
DB 223 QMRPRRLHRSVSKRWKCELTUVADAKVETHGQPVESYVLTIMNMVAGLFDHPS 282
QY 410 MGK-----DGT-----RHDAHLLLT 424
DB 283 IGNPHTITVRLVLEDEEDLKITHADNTLKSFCWKQKSNMKGDAHPHJHDAITLLT 342
QY 425 GLDICSWNKPCDTIGFAPISGCKYRSCTINEDTGLIATIAHSGHNGMHDGEG 484
DB 343 ROLCHANNRPCTEGLSHVAGNCPHRSCTINEDTGLIATIAHSGHNGMHDGEG 402
QY 485 NKCKK--SEGNTWSTPLAGNNGVSNFSCSRVYKHLKSLTAQAICLAQDP-PKPYKYP 541
DB 403 NUCEPVGRKFPFINSOLLIDAPLHNSCRQYITRFLDKRGWGLCLDUDPAKIDID--FP 460
QY 542 EXLPBELYDANTQCKWQGEKAKLMLDFFKOICAKLMCHRIGRKCTEYFPAABOTICG 601
DB 461 SVFPGGLYVSHQCLQYGLYSAFPC--EDMDNVCHTLWC--SVGTTCKSLDAANDGTRCG 517
QY 602 HDMKRGQCYKQDEOPKETHGNSWMSWQSCPCGCGHSHSRCLCTNPKSHGKCP 661
DB 518 ENKACLSEGCVPVGRF-PEAVDGSWSWSAWSTCSRSGCVGVSARHCCCTQPTPKYKRY 576
QY 662 CEGSTRTLKLCNSQKCPDSVDFRAAOCAEHNSRFRGRHYKPKYTOVEDODLCKLYCI 721
DB 577 CVGERKFRLLCNLAQCAPRSPFRHVQCSHFDAMLYKGLHTWVPV--VNDVNPCELHCR 634
QY 722 AEGFDFFFSLSNVKADGTPCSE--DSRNVICDIGERVCNDVLGSDAVEDYCGVGNH 779
DB 635 PANEFYAKLRDADCVDTGTPCYQVRASRDLCINGICKNVGCDPEIDSGAMEDRCGVCHNG 694
QY 780 SACTIHRGLYTKHHHTNQYVHVMTIPSGARSIRIYENVNSTYSIVSNALRYVINGHT 838
DB 695 STCHTVSGTTEAEGLG-YVDVGLIPAGAREIRIQEVAEAAFLAKSEDPKPYFLNGW 753
QY 839 TVDMPGRYKFGSTTDFYRRSYNEPENLIATGPTNE-----873
DB 754 TIQWNGDYQVAGTTFTVARRGNW-ENLTSQGTPEPVIQVPSARGFGGSGRGVPRST 812
QY 874 -----873
DB 813 LHRGRPGVSGPSGVTEFRSEPPFAAASVTSVPSLKWPLVAHVHRRGQWQAPLGLGW 872
QY 874 -----TLIVELLQGRNFGVAVYSNPRL--GTEKOPPAQPSYTAIVR-VNSECSVS 921
DB 873 RRHLVLMGPRPLTOLLQESNFGVHYEYTHREAGGHDEVP--PVFSMHYGPWTKCTVT 930
QY 922 CGGGR 926

```

```
Db 931 CGRG 934
RESULT 5
US-09-369-364A-17
; Sequence 17, Application US/09369364A
; Patent No. 6391610
; GENERAL INFORMATION:
; APPLICANT: Apte, Suneel
; APPLICANT: Hirskainen, Tiina L.
; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
; FILE REFERENCE: 26473/4007/10-30-00
; CURRENT APPLICATION NUMBER: US/09/369,364A
; CURRENT FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 17
; LENGTH: 1081
; TYPE: PRT
; ORGANISM: Homo sapiens ADMTS-10
US-09-369-364A-17
Query Match 20.2%; Score 1057.5; DB 4; Length 1081;
Best Local Similarity 30.2%; Pred. No. 1.3e-85;
Matches 265; Conservative 120; Mismatches 286; Indels 207; Gaps 26;
QY 164 SLHLKGG-PRHDEH-----MOLRTSSSLVAPGFTVOTLKGTKSV 204
Db 35 SRL-LPGAPRGGHRSVPLLOSGLASTHFLNLTSSRLGGRVSVYTRGLAW 93
QY 205 OTLPPEDFCVQSLSRHSN-----PSHGKCEBSTR 239
Db 94 RAARPH-CLYAGHLCQASSSVAISTCGLHGLIVADEBEVLPELHGSP--KGS--- 146
QY 240 LKLCNSQKCPDS-----VDRAA-----QCAEHSRFRGHWKPKYTOVEDLUCK 287
Db 147 -----RSFESGPHCVYRSLSRHPHLOTACGVDRDEKPNKGRFWLTKPPAR--- 196
QY 288 LYCIAGDFDFSLNKKVGTGTCSEDSRNVCIIGLCISLVSTSAHMPQPKEDLTLP 347
Db 197 -----PLNTERGGQ-----GL----- 209
QY 348 DEYKSLRHYSLLSHNELNVLTVYVDKMKQKHENITTYVLTILNVSLPFD 407
Db 210 -----KSVSR-----ERYVEINVDKRWAYHGRDVEQVLAIMNVLAKLFQD 255
QY 408 GLWG-----KDG-T 415
Db 256 SLSGSTNVLVTELLLTEDOPTLEITHAGKLSDFCKWKSIVNHSCHGNAIPENVA 315
QY 416 RHDAHLLTGIDICSWKNECDTLGPAISCKSVKSCINEDTGLGIAFTAHESGHN 475
Db 316 NHDVAVLTIDYDIQYKKECGTLGLARWA-ECVSAREAAASWRTLAATSVHCHGHT 374
QY 476 FGNHIDEGNKC---KASEGNTSPETLAGRGVFSKSCRYOLHKLSTAQACLADOP 532
Db 375 FGNHIDGVGNSCGANGODPAKLAHITMKNPFWVWSCRDYITSLDGLGLCANRP 434
QY 533 KPVKEVYKPELPGELDANTCKQKGEKALCMDFKDKCKALCHIRGKCEKFM 592
Db 435 -PRQDFVPTVAFQAYDADQCRQGHVSKCKYG---EVSSELMLCKSNKICINSI 490
QY 593 PRAEETLC-GHDM---KRCGGCVKYGDGKPEFTHGHSWSSNSPSCRTCCGVSHSR 648
Db 491 PRAEETLCQTHIDKGCYKVCVPPFSR--PEGVDGANGPFWPGDSCRTCCGVSSSR 549
QY 649 LCTNKESHGKFCGEGSTRTKLQNSQCRPSVDPAACAHSHRSRFRGHWKPYT 708
Db 550 HCDSPRTIGKYCLGERRRHSNTDDCPFSQDFREVCQAFDSIPFRGKPYKWTY 608
QY 709 QVEDQDLCKLYCIAGDFDFSLSNKKVGTGTCSEDSRNVCIIGLCISLVSTSAHMPQPKEDLTLP 768
```

```
Db 609 RGGGVKACSLTSLAEGNFYTERAAAVVDGTCPCRPDVTDIVCSGCKHVGCDVLGSDLR 668
QY 769 EDVCGVCGNNSACTIHRGLYTKHHHTNQYHYHMTIPSGARSIRIYENMVSTSYISVRNA 828
Db 669 EDKCRVCGGSGACETIEGVFPSPAGAGYEDVWIPKGSVHIFIQDNLNLSHLAKGD 728
QY 829 LRYYLNGHWTVDWGRYKPSGTTFDVRRSYNEPENLIATGPTNETLIVELLFOQRNPGV 888
Db 729 QESLLLEGLFGTPQPHRLPLAGTTTQLRQGPQVQSLEALGPINASLIVMLVLAETELPAL 788
QY 889 ANEYSMPRLGTCKPPAQPSYTWAIVR-SECVSVCGG 925
Db 789 RYRFENAC-IAROSLPP-----YSNHYAPWTKCSAQCA 821
RESULT 6
US-09-369-364A-13
; Sequence 13, Application US/09369364A
; Patent No. 6391610
; GENERAL INFORMATION:
; APPLICANT: Apte, Suneel
; APPLICANT: Hirskainen, Tiina L.
; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
; FILE REFERENCE: 26473/4007/10-30-00
; CURRENT APPLICATION NUMBER: US/09/369,364A
; CURRENT FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 13
; LENGTH: 1882
; TYPE: PRT
; ORGANISM: Homo sapiens ADMTS-9
; NAME/KEY: MOD RES
; LOCATION: (468)
; OTHER INFORMATION: Xaa = C
; NAME/KEY: MOD RES
; LOCATION: (521)
; OTHER INFORMATION: Xaa = Y
US-09-369-364A-13
Query Match 20.2%; Score 1056; DB 4; Length 1882;
Best Local Similarity 28.2%; Pred. No. 4.3e-85;
Matches 278; Conservative 142; Mismatches 315; Indels 252; Gaps 33;
QY 100 RMDAEAGNHSQNTGTENOTLH-----VLTOYDLVSAVEVDHRCGVDSHEIMHQ 151
Db 17 RDLAEMG-----SPDAAAANKDLRHPROVKLETTSEYVSPIRVNALGSPPTNV-HFK 72
QY 152 RRRRAV-----AVSEVSESLHLKGPRIHDFHMDLRTSSSLVAPGFTVOTLG 197
Db 73 RTRRSINATDPWPAFASSSSSTSSQAHVRLSAGQQFLNLTANAGTAPLFTVLLG 132
QY 198 KTGKTSVOTLPED-----FCFY-----QGSLSRH-RNSPSHGKFCGEGSTRTKLCSQK 248
Db 133 TEGVNOTKYFYSEAEALKHCFYKLCQYQYLAHGRHQP-----LLRNEHK- 177
QY 249 PRDSVDFRAAQCAEHSRRFRGRHYKWKPYTOVEADLCKLYCIAEGPDFFSLSNKVKOG 308
Db 178 -----NRHKKDKKTRARKWGERINLAGDVAALN----- 206
QY 309 TFCSEDSRNVCIIGLCISLVSTSAHMPQPKEDLTLPDEYKSLRHKHRSLLASHR--N 366
Db 207 -----SGLATEAFSAYGNKTD-----NTRKTRHRTKRFLS 238
QY 367 BELNVELTVYVDKMKQKHENITTYVLTILNVSLVAPGFTVOTLG 412
Db 239 YREFEVLVADNRWVSYHG-ENLOHYILTLMISIVASIKPDSICNLINIVNLVIH 297
QY 413 -----DGTREDHAILLTGLDICSWKNEPDCDITLGPAPIS 445
```

Db 298 EQGSPISFNAQTLLKNFCQWQSHNSFGIHHDTAVLLTQDIDCR-AHQKCDITLGLAELG 356
 QY 446 GMSKYRSCITNEDTGLGLAFTTAHESGHNFMIHDEGNMCK-...KSEGNIMSPTLAG 504
 Db 357 TICDPYRSCISSEDSGLSTAFTHAELGHVFNPHD-DNNKCKEKGVSQHVMAPTLNF 415
 QY 502 RNVGWSFPCSRQYLHKFLSTAAQACLAODQPKVKEYKPEKLPGLDYDANTOCKWQGE 561
 Db 416 YTNEMWNSKSRKYITEFLODYGECLLNEPAS-RTPPLSQLPGLLYNNVKNKCELIPOGSO 474
 QY 562 KAKLWLFKDKICAKLWCHKIG-...RKCEKEMPAABGTICCHDMWCRGGQCVKYGDEG 618
 Db 475 GSQVCFYMWQ-...CRLLMNNVNDCAHKCKTQHTPWADGTECEPGKHKCFGVKEMEGP-A 531
 QY 619 PKFTGHWSMWSFPCSRCTGGGVSHRSRLCTNPKFSGHGFCEGSTRTLKLCNSQKCP 678
 Db 532 P-VTDGWSGNSPFGTCSRTCGGKTAIRSCNRPPEPKNGKVCYGRMRKFKSCNTEPCL 590
 QY 679 RDSVDFRAOQCAEHNSRRFRGHY-...KMKP-YTOVEDQDLCKLYIAGBDF 728
 Db 591 KQXDFRDEQCA-...HFDGKHFNINGLLPSVMVFPKYSGLMKDRCKLFCRVAGNTAY 645
 QY 729 FSLSNKVDGTPCSEDSRNVICDICERVGDNDVLGSDAIVEDVCGVCGNNSACTIHRGL 788
 Db 646 YQURDEVDTGTCQGDNDICVQGLCRQAGCDHLSNKKARRDKCGVCGGNSCKTVAGT 705
 QY 789 YTKHHTNQYHMYTTPSGARSIRIYEMNVS-...TSYISVNNALRRYYLNGHWTVDMP 843
 Db 706 FNTVHY--GYNTVVRIPAGATNIDVRQHSFSGTDDNYSKSGEFLNGFNFTVMA 763
 QY 844 GR-YKFSGTTDFYRSYNPEPNLIATGPTNETLIVELLFQGR--NPGVAMEYSMPRLGTE 900
 Db 764 KBEIRIGNAVVEYSGETAVERINSTRIEQLLQVLSVGLYNPDRYSFNP-...IE 820
 QY 901 KQPP-...AOPSY-... 909
 Db 821 DKQFYVNSHGWACSKPCQGRKRLVCTRSDQITVSQDRCRLPQPHITEPGT 880
 QY 910 ----TWAVY-RSRSUSG-GRCLPV 930
 Db 881 GCDLRHVASRSCENQGLGYTLDI 907

RESULT 7

US-09-369-364A-15
 ; Sequence 15 Application US/0369364A
 ; Patent No. 6392110
 ; GENERAL INFORMATION:
 ; APPLICANT: Apte, Suneel
 ; APPLICANT: Hirsakainen, Tiina L.
 ; APPLICANT: Hirohata, Satoshi
 ; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
 ; FILE REFERENCE: 26473/4007/10-30-00
 ; CURRENT APPLICATION NUMBER: US/09/369,364A
 ; CURRENT FILING DATE: 1999-08-06
 ; NUMBER OF SEQ ID NOS: 31
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 15
 ; LENGTH: 874
 ; TYPE: PRT
 ; ORGANISM: Mus musculus ADAMTS-9
 US-09-369-364A-15

Query Match 18.9%; Score 991; DB 4; Length 874;
 Best Local Similarity 34.2%; Pred. No. 8,7e-80;
 Matches 215; Conservative 107; Mismatches 210; Indels 96; Gaps 20;
 QY 371 VETLVVVKQWQKHGHENITTYLTILNNVYSALFKDGLMK-... 412
 Db 134 VEVNVADHRVLYHG-ANLCHVLTPLMSIVASIKYDSSIGNLIVNVLWVHNEOG 192
 QY 413 -----DGTDRHAILTGLDICSWKNEPQDITGFAPIQWC 448

Db 193 PYINENQAQTTLLKNFCQWQSHNSFGIHHDTAVLLTQDIDCR-AHQKCDITLGLAELG 251
 QY 449 SKYRSCITNEDTGLGLAFTTAHESGHNFMIHDEGNMCK-...KSEGNIMSPTLAGRNG 504
 Db 252 DPYRSCISSEDSGLSTAFTHAELGHVFNPHD-DNNKCKEKGVSQHVMAPTLNFVTN 310
 QY 505 VFNWSFPCSRQYLHKFLSTAAQACLAODQPKVKEYKPEKLPGLDYDANTOCKWQGEKAK 564
 Db 311 PMWNSKSRKYITEFLODYGECLLNEPAS-RTPPLSQLPGLLYNNVKNKCELIPOGSO 369
 QY 565 LCMLEDFKDKICAKLWCHKIG-...RKCEKEMPAABGTICCHDMWCRGGQCVKYGDEGKRP 621
 Db 370 VCFYMWQ-...CRLLMNNVNDCAHKCKTQHTPWADGTECEPGKHKCFGVKEMEGP-A 425
 QY 622 THGHWSMWSFPCSRCTGGGVSHRSRLCTNPKFSGHGFCEGSTRTLKLCNSQKCP 681
 Db 426 IDSWSGNSPFGTCSRTCGGKTAIRSCNRPPEPKNGKVCYGRMRKFKSCNTEPCKMQK 485
 QY 682 VDFRAOQCAEHNSRRFRGHY-...KMKP-YTOVEDQDLCKLYIAGBDF 731
 Db 486 RDEFREQCA-...HFDGKHFNINGLLPSVMVFPKYSGLMKDRCKLFCRVAGNTAYQOL 540
 QY 732 SNKYVDGTPCSEDSRNVICDICERVGDNDVLGSDAIVEDVCGVCGNNSACTIHRGLYTK 791
 Db 541 RDRVIDGTGTCQGDNDICVQGLCRQAGCDHLSNKKARRDKCGVCGGNSCKTVAGT 600
 QY 792 YTKHHTNQYHMYTTPSGARSIRIYEMNVS-...TSYISVNNALRRYYLNGHWTVDMP 845
 Db 601 VHY--GYNTVVRIPAGATNIDVRQHSFSGTDDNYSKSGEFLNGFNFTVMA 658
 QY 846 YKFSGTTDFYRSYNPEPNLIATGPTNETLIVELLFQGR--NPGVAMEYSMPRLGTEKOP 903
 Db 659 VRGSAVIRYSGSDNVCRLNCTRIEQLLQVLSVGLYNPDRYSFNP-...IEDKP 715
 QY 904 KQPP-...TWAVY-RSRSUSG-GRCLPV 926
 Db 716 ----QFTVNSHGWACSKPCQGRKRLVCTRSDQITVSQDRCRLPQPHITEPGT 737

RESULT 8

US-09-130-491-2
 ; Sequence 2 Application US/09130491
 ; Patent No. 6416974
 ; GENERAL INFORMATION:
 ; APPLICANT: Holtzman, Douglas A.
 ; APPLICANT: Holtzman, Douglas D.
 ; TITLE OF INVENTION: TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83
 ; FILE REFERENCE: 09404/041001
 ; CURRENT APPLICATION NUMBER: US/09/130,491
 ; CURRENT FILING DATE: 1998-08-07
 ; EARLIER APPLICATION NUMBER: US 60/058,108
 ; EARLIER FILING DATE: 1997-09-05
 ; EARLIER APPLICATION NUMBER: US 60/054,961
 ; EARLIER FILING DATE: 1997-08-06
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 2
 ; LENGTH: 967
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-130-491-2

Query Match 18.8%; Score 984.5; DB 4; Length 967;
 Best Local Similarity 30.9%; Pred. No. 3.9e-79;
 Matches 270; Conservative 104; Mismatches 328; Indels 173; Gaps 35;
 QY 157 VAVSEVE-...SLHLRLKGRPHDMLDRTSSLSVAPGPIVQLG-KTGTSGVQLPP 209
 Db 60 LVWPELRAPGHGTTLRLHAFQQLDLELAPDSSFLAPGTLQNVGRKSSGT-...PLPE 117
 QY 210 EDF-...CFYQGLSHRNSFSGCKFCGEGSTRTLKLCNS-...QKCPDSVDFRAOQCAEH 263
 Db 118 TDLAHCYSGTVNGDPSSAA-ALSLCEGVRGAFVLLGAYFIQPLPAASERLATAPE- 175

QY 264 NSRRFRGHKKYKPTQVADLCKLYCIARGDFPFSSLNKVKDGTSCSDSRNNVCDIGI 323
 DB 176 -----XPPAPLQ-----FLLRRNRQD-----VGGT 197
 QY 324 CELSVSTSAH-----MPQPKEDLPILDEYKSLPKHKSIL 360
 DB 198 C-GVVDPRPGKAETDEDECTEGEDGQWSPQALQGVQPGTGS-IRKX-F 253
 QY 361 LSPHSEBELNVLTVVVKMKMONGHENITTVLTILNVSALPK----- 406
 DB 254 VSSIR---YVETMLVADQSKAEFFG-SGLKHYLLTFLSVAARLYKPSIRNSVSLVVK 338
 QY 407 -----DGLMG-----XDGTRHDHAILLTGLDICSWKNEP 435
 DB 309 ILVIHDEKGFVETNSAALTLRNFNCWKQKHNPSPDRDAHYDTAILTRQDLG--SQT 366
 QY 436 CDTLGFAPISGMSKYRSCITNEDTGLGTAFTAHESGHNFHIDGEGNMCKSEG--- 492
 DB 367 CDTLGMADVCTGCPERSCSVIEDDGLQAFTAHELGHVFNPHD-DANQCSASLVNQ 425
 QY 493 --NIMSPITLAGRVFSPSCSYQIHLKFLSTAAICLADQPKPKVKEYPEKLPGLGYD 550
 DB 426 DSHMASMLNLDHSPMSPCSAYMITSFLDNGHGECLMD--KPQNPILQPLGDLGTSYD 483
 QY 551 ANTCQKQWQGEKAKLCLDFKDKICAKLWCHRIGRK-----CETKFMFAEAGTICGHDMW 605
 DB 484 ANRQQTFTGDSKHC--PDAASTCTTLWC--TGTSGGLLVCTQKHFPMADGTSCEGKW 539
 QY 606 CRGQCVAYGDEG--PKFTGHMSDWSWSPSCRTCGGVSYSRSLCTNPKFSGHGRFCE 663
 DB 540 CINGCVNKTDRKHFDTPFHGSMGWGPDGDSRTCGGQVQYTMRECDNPPVKNKGKCYE 599
 QY 664 GSTRTLKLNQKCP-RDSVDFAAQAEBN---SRRFRGRH--YKWKP-YTOVEDQDLCK 717
 DB 600 GREVYSRNLCDPNNNGKTFPEQCAHNEFSKASGPGPAVETPKTAGVSPKDRCK 659
 QY 718 LYCIABGDFPFFSLSNKVKDGTSCSDSRNNVCDIGICERVCNVLGSDADEVGVGNG 777
 DB 660 LIOAKGIGTFVLPQPKVDGTSCSDSTVSVQGVQKACDRIIDSKKKFKDGVGCG 719
 QY 778 NNSACTIHRGLYKTHHNTQYHMTIPSGARSIRIYEMV-----STSYISVRLARY 832
 DB 720 NGSTCKIJSVTS--AKPGYHDIITIPGATNIEVKQRNQRNNGSFLAKAADGT 777
 QY 833 YLNGHWTVDPGR-YKSGTTFYRSYNEPENLIATGPNETLIVELLFQGR--NPGVA 889
 DB 778 ILNGDTLSTLDQIMYGVLYSGSAALERSFSPLEPLTIQVLMVGHALRPKIK 837
 QY 890 WYSMERLCTGKOPAPQSYT-WAIVR-SECSVSC 922
 DB 838 YTFYVK--KKEFNAIPTSEWVIEWGECSKSC 870

RESULT 9

US-09-445-023A-12
 ; Sequence 12, Application US/09445023A
 ; Patent No. 656589
 ; GENERAL INFORMATION:
 ; APPLICANT: Hirose, Kunitaka
 ; APPLICANT: Inoguchi, Biiji
 ; APPLICANT: Hakezaki, Michinori
 ; APPLICANT: Ishikawa, Keiko
 ; APPLICANT: Ishida, Yukako
 ; APPLICANT: Matsushima, Kouji
 ; APPLICANT: Kuno, Kouji
 ; TITLE OF INVENTION: Human ADAMTS-1 protein, gene encoding the same, pharmaceutical
 ; FILE OF INVENTION: composition and method of immunologically analyzing human ADAMTS
 ; FILE REFERENCE: Q57092
 ; CURRENT FILING DATE: 1999-12-03
 ; PRIOR FILING DATE: JP 9-160422
 ; PRIOR FILING DATE: 1997-06-03

; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: Patent in version 3.0
 ; SEQ ID NO: 12
 ; LENGTH: 727
 ; TYPE: PRT
 ; ORGANISM: Mus sp.
 ; US-09-445-023A-12

Query Match

Best Local Similarity 35.3%; Pred No 5 3e-72;
 Matches 222; Conservative 77; Mismatches 241; Indels 89; Gaps 22;

QY 371 VETLVVVKMKMONGHENITTVLTILNVSALPK----- 406
 DB 20 VETMLVADQSKAEFFG-SGLKHYLLTFLSVAARLYKPSIRNSISLVVKILVIEBQK 78
 QY 407 -----DGLMGKDGTRHDHAILLTGLDICSWKNEPCDTLGFAPIS 445
 DB 79 PEVTSNAALTLRNFNCWKQKHNPSPDRDAHYDTAILTRQDLG--SHTCTLGMADV 136
 QY 446 GMSKYRSCITNEDTGLGTAFTAHESGHNFHIDGEGNMCKSEG-----NIMSPITLA 500
 DB 137 TVCDPFSRCSVIEDDGLQAFTAHELGHVFNPHD-DAKHCSASLVNGSDSHMASMLS 195
 QY 501 GRNGVFSWSPSCSYQIHLKFLSTAAICLADQPKPKVKEYPEKLPGLGYDANTCQKQW 560
 DB 196 SLDSHSPMSPCSAYMITSFLDNGHGECLMD--KPQNPILQPLGDLGTLGDANRQCF 253
 QY 561 EKAKLCLDFKDKICAKLWCHRIGRK-----CETKFMFAEAGTICGHDMWCRGQCV 615
 DB 254 EESKHC--PDAASTCTTLWC--TGTSGGLLVCTQKHFPMADGTSCEGKWCVKCN 309
 QY 616 D--EGKFTGHMSDWSWSPSCRTCGGVSYSRSLCTNPKFSGHGRFCEGSTRTLKLN 673
 DB 310 DMKHFATPHGSMGWGPDGDSRTCGGQVQYTMRECDNPPVKNKGKCYEKGKVRYS 369
 QY 674 SQKCP-RDSVDFAAQAEBN---SRRFRGRH--YKWKP-YTOVEDQDLCKLYCIAEG 727
 DB 370 IEDCPDNNNGKTFPEQCAHNEFSKASGPGPAVETPKTAGVSPKDRCKLTCEAKG 429
 QY 728 FFLSNKVKDGTSCSDSRNNVCDIGICERVCNVLGSDADEVGVGNGNNGSACTIHRG 787
 DB 430 FVLQPKVDGTSCSDSTVSVQGVQKACDRIIDSKKKFKDGVGNGSTCKKMSG 489
 QY 788 LYTKHHNTQYH-MTIPSGARSIRIYEMV-----STSYISVRLARYYLNGHWTV 841
 DB 490 IVTS---TRPGYHDIITIPGATNIEVKQRNQRNNGSFLAKAADGTILNGNFTLS 546
 QY 842 WGR-YKSGTTFYRSYNEPENLIATGPNETLIVELLFQGR--NPGVAWYSMPRLG 898
 DB 547 TLEQDTYKGVLYSGSAALERSFSPLEPLTIQVLMVGHALRPKIKFTYFWK-- 604
 QY 899 TEQKPAQSYT-WAIVR-SECSVSC 925
 DB 605 KTESFNAIPTSEWVIEWGECSKTCGSG 633

RESULT 10

US-09-445-023A-1
 ; Sequence 1, Application US/09445023A
 ; Patent No. 656589
 ; GENERAL INFORMATION:
 ; APPLICANT: Hirose, Kunitaka
 ; APPLICANT: Inoguchi, Biiji
 ; APPLICANT: Hakezaki, Michinori
 ; APPLICANT: Ishikawa, Keiko
 ; APPLICANT: Ishida, Yukako
 ; APPLICANT: Matsushima, Kouji
 ; APPLICANT: Kuno, Kouji
 ; TITLE OF INVENTION: Human ADAMTS-1 protein, gene encoding the same, pharmaceutical
 ; FILE OF INVENTION: composition and method of immunologically analyzing human ADAM
 ; FILE REFERENCE: Q57092
 ; CURRENT APPLICATION NUMBER: US/09445,023A

; CURRENT FILING DATE: 1999-12-03
 ; PRIOR APPLICATION NUMBER: JP 9-160422
 ; PRIOR FILING DATE: 1997-06-03
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 1
 ; LENGTH: 727
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-445-023A-1

Query Match 17.8%; Score 934; DB 4; Length 727;
 Best Local Similarity 34.4%; Pred. No. 8.4e-75;
 Matches 221; Conservative 82; Mismatches 247; Indels 92; Gaps 22;
 QY 354 LRKSLRLSHRNEELNVLVVDKMMQNHENITVYLILNMVSALEK-----406
 DB 8 LRKRFV-----SFRPYETNLVADQSNABFHG-SGLKHLVLLFSAALRYKHSIRNS 61
 QY 407 -----DLWG-----KQTRDHAILTLGLDI 428
 DB 62 VSLVVKVILVDEQKEVEVTSNALIRLPCWQKQNPSPDRDAEYDITALLTQDL 121
 QY 429 CSWKIEPCDPLGAPISCMKYSRCTINEDTGLAFTTAHSGHFGMHDGKNCK 488
 DB 122 CG--SQTUHLGADVGVCDPSRCSVEDDGLQAFTTRELGHVFNWPHD-DAKQCA 178
 QY 489 KSEB-----NIMSPFLAGNGVSPSCRSQYLKFTSTAQICLADQPKYKVEK 543
 DB 179 SLNGVQDSHMMASLMSLDSOPHSPCSAYMITFSLNGHGECLMD--KEQNPQLQPD 236
 QY 544 LPEGLYDANTCKWQFGEKAKLMLDFKDIKCALWCHRTGRK-----CETVEMAAEGT 598
 DB 237 LPTGLYDNRCCQTFEGDSKHC--PDAASTCTSLWC--IGTSGGVLCVCTHFFWADET 292
 QY 599 ICHDWMCWGSCQCVKYGEG--PKPCHGHSDWSSMSPCRCTCGGYSRSLCTNPKPS 656
 DB 293 SCCEGWGICNGKWKTRKHFTDFPHSGMPGMPGWCDCSTCGGQVQVYTWRECDNPVK 352
 QY 657 HGKFCBGSRTLLKLCNSQKCP--RQSVDFRAAQCAEHN---SRERFGRHYKWK--YQV 710
 DB 353 NGKCYCEGRVRYRSNLEDCPDNNGKTFRECEAHNEFSKASFGSPAVEMIPKAGV 412
 QY 711 EDQDLKLYCTAEGDFFFSLSNKKVDGTPCSDSRNVCIDGICERVCNDVLGDAVED 770
 DB 413 SPDRCKLIQKANGIGYFVLQPKVVDGTPCSPDSTSVQGVQCVKACDRILDSKKFD 472
 QY 771 VCGVGNNSACTIHRGLYTKHHTNQYTHMVTIPSGARSIRIYEMNV-----STSYLS 825
 DB 473 KCGVCGNGSGCKKISGSVTE--AKPGYHDIPTGATNIEVKORNGSRNCSFLAI 530
 QY 826 RNALRYLNGHWTVDWPCR--YKFSGTTFDYRSYNEPENLIATGTFNTEILVELLFGGR 884
 DB 531 KAADGTYLNDGYTLSTLEODIMYKWLRYSGSNALEIRSFSLKEPLATLOVLTGVN 590
 QY 885 --NPGVAMEXMPRLGTEKQPPAQPSYT--NAIVR--SECVSC 922
 DB 591 ALRPKIKYTVFKK--KKESFNAIPTFSANVIEWGECSKSC 630

RESULT 11
 US-09-491-522-11
 ; Sequence 11, Application US/09491522
 ; Patent No. 6428998
 ; GENERAL INFORMATION:
 ; APPLICANT: Colige, Alain
 ; APPLICANT: Lapierre, Charles M.
 ; APPLICANT: Prockop, Darwin J.
 ; TITLE OF INVENTION: RECOMBINANT N-PROTEINASE,
 ; TITLE OF INVENTION: AND THE PRODUCTION, METHODS AND USES THEREOF
 ; NUMBER OF SEQUENCES: 17
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds, LLP

; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: NY
 ; COUNTRY: USA
 ; ZIP: 10036-2811
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: Windows
 ; SOFTWARE: FastSeq for Windows version 2.0b
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/491,522
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/886,333
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Abrams, Samuel B
 ; REGISTRATION NUMBER: 30,605
 ; REFERENCE/DOCKET NUMBER: 9389-0060-999
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 650-493-4935
 ; TELEFAX: 650-493-5556
 ; TELEX: 66141 PENNIE
 ; INFORMATION FOR SEQ ID NO: 11:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1205 amino acids
 ; TYPE: amino acid
 ; STRUNDESS: single
 ; TOPOLOGY: linear
 US-09-491-522-11

Query Match 17.8%; Score 933; DB 4; Length 1205;
 Best Local Similarity 28.9%; Pred. No. 2.1e-74;
 Matches 277; Conservative 125; Mismatches 356; Indels 200; Gaps 34;
 QY 43 ASPP-RLLSRPERLTAMSPLSA-GTCVRHGTGSGSAWEPPASSSTRCAA-----GL 95
 DB 35 ADFFGQPGHGAERILAVPRTDAQRLVSHVSAATA----PAGVTRRAAPAQIPGL 89
 QY 96 DGKGRMDEACNHRSSOOTNTGTENQTLHLVQYDLVSAVEVDHRGDYVSHHQRERR 155
 DB 90 SG-GSEDFGCR-----LFVNYTVFGR-----110
 QY 156 AVAVSEVESLHLKGPGRDFHMDLTSSLVAPGFIQTLGTGTSVOTLPEDPCFY 215
 DB 111 -----DLHLRLAFNARLVAFGATVEMQGESGATRVE--PLLCTCLY 149
 QY 216 QGSLRSHRNSPHGKFCBGTSTTLKLCNSQKCPDRSDVDFRAAQCAEHNRRFRGR-HYK 274
 DB 150 VGDVAGLAESSVALSCNDGLAGLIRNMBEEFFIPELEKGLAAKEA-----QGRVHV 203
 QY 275 WKPYTOVEDADLCKLYCIAEGDFFFSLSNKKVDGTPCSDSRNVCIDGICELSVSTSAH 334
 DB 294 YHRAFTTSR-----PFFLGGFQALD--TGISADS-----LDGL-----233
 QY 335 MPQPPKEDLFLPDEYKSLCHRSRLSHRNEELNVLVVDKMMQNHENITTV 394
 DB 234 -----SRALGVLEERVNSRRMR---RHAADDDYNIIEVLVLDVDSVQPHGTHVOKYL 285
 QY 395 LTIANNVSALEKPDGLMG-----K 412
 DB 286 LTLANIVNEIYHDESIGAHINVLVRIILSYGKSMLEIGNPSQSLNVCNWAYLOQK 345
 QY 413 DGT----RHDHAILTGLDICSWKNSPCDTLGFAPISCMKYSRCTINEDTGLGIATTI 468
 DB 346 PDTDDHVDHDAIFLTLRODF-----GPSGMQGYAPVTGMCHPVRSCTUNHEDFSSAFV 400
 QY 469 AHESGHNFQMTHDGONWC--KXSEGNITSPFLAGNGVSPSCRSQYLKFTLSTAQAI 526
 DB 401 AHETGVLGMEHDGQNRCCGDEVRLGSLNAPLQQAFAHFRHNSRCSQQLSYHSYD--458

QY 527 CLADOPKVEKYPEKLGELYDANTOCKWOFGEKAKLMLDFKDI CKALWCHRTGRK 586
 Db 459 CLRDDPF-THDPALPQLPBLHYSNNEQRFDFGLGYMCTAFRTDFPCKQLWCHSDNP 517
 QY 587 --CETKMPAAECTTICGHDMWCEGCKVKGDEGKPTGHGHSWMSWSPCSRTCCGGVS 644
 Db 518 YFCKTKGFLDGTWCAPGKHCFKHCTMTLTDILK-RDGNWAGWSFPGSCSRTCTGTGVK 576
 QY 645 HSRSLCTNPKPSHGCKCFCESTRTILKCNCSQKPRDSVDFRAOCA-----EHSRFR 698
 Db 577 FTRQCDNPHFANGSRTCSGLAVDFQLCNSQCPDALADFREQCROWLVEHSDAQ-- 634
 QY 699 GRHYKPKFYTOVEDDOLCKYCIAGDFDFPFLSNKYKDTPCS-EDSRNVICIDICERV 757
 Db 635 ---HHMLPHBRDAKERCHLYCESKGTGEVVSUKGMVHDTRCSYKDAFSLCVRGCRKV 691
 QY 758 GCDNVLSGDAVEDVCGVCGNNSACTHRLGYTKHHTNOYVHMVTIPSGARSIRIYEMN 817
 Db 692 GCDGVIGSKQEDKCGVCGGDNCHCKVVKGTESFKLGYIKMFEIPAGARHLILOAD 751
 QY 818 VSTSIVSVN-ALRYLYNGHMTVDMWRKYF--SGTTFYRSYNEFENLATGPTNET 874
 Db 752 TTSHLAVKMLTGKFLNEENDVD-PNSKTFIANGVEMEY-RDEGRETLOTMPLRGT 809
 QY 875 LIVELLFQG-RNPGVAMEYSM--PRLGTGRQPPAQF---SYTMAIVR-SECSVSCGG 925
 Db 810 TTVLVEIGDARISITYKMIHEDSLNVDDNVDSDSVGYEWALKKWSFKPCGGG 867

RESULT 12
 US-09-491-522-5
 ; Sequence 5, Application US/09491522
 ; Patent No. 6428998
 ; GENERAL INFORMATION:
 ; APPLICANT: College, Alain
 ; APPLICANT: Lapiere, Charles M.
 ; APPLICANT: Prockop, Darwin J.
 ; TITLE OF INVENTION: RECOMBINANT N-PROTEINASE,
 ; TITLE OF INVENTION: AND THE PRODUCTION, METHODS AND USES THEREOF
 ; NUMBER OF SEQUENCES: 17
 ; CORRESPONDENCE ADDRESSES:
 ; ADDRESSES: Pennie & Edmonds, LLP
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: NY
 ; COUNTRY: USA
 ; TELEPHONE: 212-693-8811
 ; FAX: 212-693-8811
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: Windows
 ; SOFTWARE: Fast-Seq for Windows Version 2.0b
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/491,522
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/886,333
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Abrams, Samuel B.
 ; REGISTRATION NUMBER: 30,605
 ; REFERENCE/DOCKET NUMBER: 8389-0060-999
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 650-493-4935
 ; TELEFAX: 650-493-8556
 ; TELEX: 66141 PENNIE
 ; INFORMATION FOR SEQ ID NO: 5:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1211 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear

US-09-491-522-5
 Query Match 17.6%; Score 923; DB 4; Length 1211;
 Best Local Similarity 28.6%; Pred No. 1.8e-73;
 Matches 266; Conservative 132; Mismatches 363; Indels 168; Gaps 31;

QY 81 PERPASSSTGAGAGLDGGRDMDENAGHRSQOITGTENTLHVLTVQVLDVSAVEVDHRG 140
 Db 29 PPPPPANARLAAAADPPGGL-----GHGAERI-----LAVPVRTDAQG 68
 QY 141 DYSVSEHIMHQ-----RRRAVAV-----SVESLHL--RLKGPRIHDMRLTSS 184
 Db 69 RLNVSHVSAATSAGVRAARAAPVPTSPFTGGNEEPEGSHLFVNVTVPGRDLHLRLDNA 128
 QY 185 SLVAPQETVOTLCKGTGKTSVQTUPPEDFCYQSLRSHRSRSHSGSKFCGSGSTRTILKCN 244
 Db 129 RLVAQATMEWQKGTTRVE--PLIGSLCYVGDVAGLAESAVALSNCDGLAGLIRME 186
 QY 245 SQKCPRDSVDFRAACACSHNSRRFRGR-HYKM-KPVTQVEADLCKLYCIAEGDFEFLS 302
 Db 187 EEFPIEPLKGLAAQEA-----QQRVHVVVYRPPPTSPPL----- 221
 QY 303 NKVKDGTTCSEDSRNVICIDICELSVSTSAHPOPPKEDLFTLPDEYKCLAHKRSLLR 362
 Db 222 ----GGPQALDT-GASLDSLSL-----RALGVLEHANSRRRAR--R 259
 QY 363 SHRNELNVETLVVVDKMKWQNHGHENITTVVITILNNVSALFKDGLAG----- 411
 Db 260 HAADDVNEIVLLGVDDSVVQFHGKEHVQKYLTLNNTIVNEIYHDESLGAHINVVLRII 319
 QY 412 -----KOGT-----RHDHAILTLGLDICSWMKEPC 436
 Db 320 LLSYKGKMSLEIGNPSQSLNVCWMAYLQOKPDTGHDYHDAIFLTRDP-----GPS 374
 QY 437 DTLGAPISGMSKVSRSCTINEDTGLGLAFTIAHESGHNEFMTHGBONMC--KKSEGI 494
 Db 375 GMDGAPVTGMCHPVRSCITLHEDGSFAFWAHETGHVGMHDGQGNCGGDEVRLSGI 434
 QY 495 MSTPLAGRNQVFSNPSQSLNHLFLSTAAQAI CLADQPKPKVBYKPKPLGELYDANTQ 554
 Db 435 MAFVQAAAFRRFHWSCSQELSYLHSD--CLLDDPF-AHDMFALPQLPGLHYSMNEQ 491
 QY 555 CKMQFGKAKLWMLDFKDI CKALWCHRTGRK--CETKMPAAEGTICGHDMWCEGCKV 612
 Db 492 CRFDGLGYMCTAFRTDFPCKQLWCHSDNPYFCKTKGPPUGTMCAPGKHCFKHCTI 551
 QY 613 KYDGGPYPFHGHSWMSWSPCSRTCCGGVSHRSRLCTNPKPSHGSKFCGSGSTRTILKLC 672
 Db 552 WLTPDLK-RDGSNGWSFGSGSRTCGTGVKTRQCDNPHFANGSRTCSGLADPOLC 610
 QY 673 NSQKPRDSVDFRAOCA-----EHSRFRGRHYKWKPKYTOVEDDOLCKLYCIAEGFD 726
 Db 611 SRQCDPSLADFRESCQWDLVFENDQ-----HHWLFARHDAKREKCHLYCSRTG 665
 QY 727 FFEISLNVKDGDTPCS-EDSRNVICIDICERCVGDNVLGSDAVGCVGNGNSACTIH 785
 Db 666 EVSMKRVWHDGTRCSYKDAFSLQVRGDKKCGGIGVSGSKQKCGVCGGNSHCKV 725
 QY 786 RGLYTKHHTNOYVHMVTIPSGARSIRIYEMNVSTYSIVN-ALRYLYNGHMTVDMWR 844
 Db 726 KGTFRSPKGGYIKWFEIPAGARHLILOEVDATSHLAVKMLTGKFLNEENDVASS 785
 QY 845 R-YKSGTGTDFRSYNEFENLATGPTNETLTVLLFQG-RNPGVAMEYSM--PRLGT 900
 Db 786 KTIANGVEMEY-RDEGRETLOTMPLRGTITVLVTPVGDTRVSLTYKMIHEDSLNVD 844
 QY 901 KQPPAQPS---YTMAIVR-SECSVSCGG 925
 Db 845 DNNVLEDSVVYEWALKKWSFKPCGGG 873

RESULT 13
 US-09-130-491-13

Sequence 13, Application US/09130491
 Patent No. 6416974
 GENERAL INFORMATION:
 APPLICANT: Holtzman, Douglas A.
 APPLICANT: Goodearl, Andrew D.J.
 TITLE OF INVENTION: TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83
 FILE REFERENCE: 09404/041001
 CURRENT APPLICATION NUMBER: US/09/130,491
 CURRENT FILING DATE: 1998-08-07
 EARLIER APPLICATION NUMBER: US 60/058,108
 EARLIER FILING DATE: 1997-09-05
 EARLIER APPLICATION NUMBER: US 60/054,961
 EARLIER FILING DATE: 1997-08-06
 NUMBER OF SEQ ID NOS: 16
 SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 13
 TYPE: PRN
 LENGTH: 608
 ORGANISM: Mus musculus
 US-09-130-491-13

Query Match 17.2%; Score 903; DB 4; Length 608;
 Best Local Similarity 36.5%; Pred. No. 3.8e-72;
 Matches 206; Conservative 73; Mismatches 236; Indels 50; Gaps 20;

QY 412 KQTRHDAIILLTGLDICSWNKPEPTGLFAPISGMSKYSRSTINBTGLGATIAHE 471
 DB 40 RDAERTRLTFRQDLCG--SGTCUOLMADGVICDPSSCSVIEDSLQAFITIAHE 97
 QY 472 SCHWFGKTHDCEGMMCKKSEC-----NIMSPTLAGNGVFSMSPGCSROYLHKLSTAGAI 526
 DB 98 LGHVFNPHD-DAKQASLNGVQDSHWMASMLNDSQSPSCSAYMITSFLDNGHGE 156
 QY 527 CLADQPKPKYKYPEKLPGLSELVDANTCKWQGEKAKLWLPDKDICKALWCHRIGRX 586
 DB 157 CLMD--KQNPQIQPCDLPSTGYDANRQCOFTGEDSKHC--PDAASTCTLAC--TGTS 210
 QY 587 -----CETKMPAAETLICHDMKGGCCVKYGDGEG--PKETHGHWSADSSMSPCSRTC 639
 DB 211 GGVLVQTKHFPWADTSGSGKWKCKNGKCNKNTDKHFDPFHSGMWGMPGDCSRTC 270
 QY 640 GGVSHRSLCTPKPSHGKFCGSTRTLKLNCKPC-RDSVDFRAAQAEHH-----S 694
 DB 271 GGGVQYTMRECNFVPMGKGYCEGKEVYFSCNLEDCEPDNCKTFREQCBAHWERSKA 330
 QY 695 RFRGRHYKWP-YTOVEDQDCLKYCIAEGFDPFSLSNKVKDGTGCSDSRNYCIDGI 753
 DB 331 SFGSGPAVEMPKYAGVSPDKCKLCOAKGIGYFVLQPKVYDGTGCPSPDSTSCVQCG 390
 QY 754 CERVGDVYLSGDAVEDCGVGNNGNSACTTHRGLVTKHHHTNQYVHVMTIPSGARSIRI 813
 DB 391 CVKAGCDRIIDSKKPKDKCGVCGNGSTCKKISGVSYS--AKPGYHDIITPIGTATNIEV 448
 QY 814 YEMNV-----STSYISVNRALRYLYNGHWTVDPOR-KYFSGTTFDYRYSNPEENLIA 867
 DB 449 KQRNORGRNNGSFLAIAADGYLYNGDYTLSTLEODIMYKGVLYRSGSSAALERTS 508
 QY 868 TGPNTETLVELLFQCR--NPGVAWEPYNSRPLETKETQPPQAPSYT-MAIVR-SECVSQC 923
 DB 509 FSPLEKPTIOLVTVGNALRPKNTYFYFKK--KESFNATPTFSAWVIEENGECSTCG 566
 QY 924 GG-----RCLP---VLLLEAACOP 939
 DB 567 GYKYSKSLKCLSHDGGVLSHESCDP 951

RESULT 14
 US-09-369-364A-2
 Sequence 2, Application US/09369364A
 Patent No. 6391610
 GENERAL INFORMATION:
 APPLICANT: Apce, Suneel
 APPLICANT: Hurskainen, Iina L.

APPLICANT: Hirohata, Satoshi
 TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
 FILE REFERENCE: 26473/4007/10-30-00
 CURRENT APPLICATION NUMBER: US/09/369,364A
 CURRENT FILING DATE: 1999-08-06
 NUMBER OF SEQ ID NOS: 31
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 2
 LENGTH: 930
 TYPE: PRN
 ORGANISM: Mus musculus ADAMTS-5
 US-09-369-364A-2

Query Match 16.9%; Score 885.5; DB 4; Length 930;
 Best Local Similarity 28.7%; Pred. No. 2.8e-70;
 Matches 273; Conservative 120; Mismatches 354; Indels 205; Gaps 40;

QY 48 LLSRGRRLTAMSPLES-AGTCVRRHGTSGSAWEPERPASSSTRAAGLKGKGRMDRAG 106
 DB 14 LLSASCLSLAADSFAAQAQDKTROPQAAAAAEPPQGEETR-----ERHLOPLAG 67
 QY 107 NHRSQTTNGTENQTLHVLQYOLVAYVDHGVSHHMHQRRRAVAVSEVSLH 166
 DB 68 QRKS-----GLVNRIDQ--LISG-----GKVGILVYAGRR----- 98
 QY 167 LRKGRHDPHMDLTSSLSVAGHGVQTKGTGKTSVQTLPPEDGFCYQSSLRHNSP 226
 DB 99 -----FLDLERDTVGAAGSIVTAGGGLSASSGH-----RGCFYRGIV----- 138
 QY 227 SHGKFCGSEFRTL--KLNSQKCPROSDVDPAAOCHAHNSR-RPRGRHYKWKYTOVE 282
 DB 139 -----DGSRLSAVFLDGLD-----QFFAVKHARYTLKPLRG--SNAEYERY 182
 QY 283 AD-----LCKLYCTAGDPFPFSLSNKVKDGTGCSDS-----RNVCIDGICE 325
 DB 183 GGGSRHLVY-NRSGFS-FEALPPRASCETPASPGSQSPSWHSRSLRSLAPQLLD 240
 QY 326 LSVVSPSAHM-POPKEDLFDLPEYKSLRHKSLSHSHENLNVTLVVYDKKMWON 384
 DB 241 HSASFSPGNAGPQ-----TWRRRRRSISBARQ-----VELLVADSSMARM 282
 QY 385 HGHENITTVTLTANMVSALP-----KD----- 407
 DB 283 YG-RGLQHYLLTWASIANRLYSHASTENHRLAVVKKVVVLTDKDTSLVSKNAATTLKNF 341
 QY 408 -----GLWKGKGTTH-DHAILLTGLDICSWNKPECDTGLCFAPISGMSKYSCTINED 459
 DB 342 CKWQHQNQLGDHHEHYDAAILFTREDLQ--HHSCTDLGMADVGTICSPERSCAVIED 399
 QY 460 TGLGLAFTIAHESCHNFNKHDEGNMCKKSEG-----NIMSFTLAGNGVFSWSPCSRO 514
 DB 400 DGLHAAFTVAHEHCHLGLSHD-DSKCEENFTGTEDKRLMSSILTSIDASKPWSKCTSA 458
 QY 515 YLHKPLSTAQAICLADQPKPKYKYSKLPGLBLYDANTCKWQGEKAKLWLPDKDI 574
 DB 459 TITFELDDHGNCCLDLPR--KQILGPBELFGQTDATQCNLTGPEYSVCP--GMDV 513
 QY 575 CKALWCH--RTGRK-CETKMPAAECTICGHDMWCGGQCVKYGDEGPK-----PTHGM 626
 DB 534 CARLWCAVVRGQMVCLTKLPAVEGTGCGKRVCLQKCV--DKTKKYYSTSHGNW 570
 QY 627 SDMSWSPCSRTCGGVSHASRLCTNPKPSHGKFCGSTRTLKLNCKSQKCPDSVDFA 686
 DB 571 GSGMPGQCSRSSCGGVQFAYRHCNNPAPNSRGYCTGCAIYRSCSVTPCPNGKSPRH 630
 QY 687 AQCAHNSRRFRGRHYK-----WKP-YTOVEDQDCLKYCIAEGFDPFSLSNKVKDGPFC 741
 DB 631 EQCEANQYQSDAKGVKTFVWPKYAGVLPADVCKLTKRAKGTGYVYVSPKVDGTGTEC 690
 QY 742 SEDSRNVCIDGI CERVGDVYLSGDAVEDCGVGNNGNSACTTHRGLVTKHHHTNQYH 801
 DB 691 RPYNSVSVGRVCRTGCTGIIIGSKLQYDKGCGVGGDNSSCTKIIGTNK--KSKGYTDV 748

QY 802 VTIPSGARSIRIYEMNVS-----TSYISVRNALRYLYNGHWTVDWPGR-YKPSGTTFDY 855
 Db 749 VRIPGATHIKVROPKAKDQTRFAYLALKKTKTGEYLINGKYMISTSETIIDINGTWNY 808
 QY 856 RRSVNEPENIATG--PTNETLIVELLFOQNP---GVAMEYSNPLRGTEK 901
 Db 809 SGWSHRDDFLHGMGYSATKEILIVQIL--ATDPTKALGVRYSFVFPKKTQK 858

RESULT 15
 US-09-122-126B-15
 ; Sequence 15, Application US/09122126B
 ; Patent No. 6451575
 ; GENERAL INFORMATION:
 ; APPLICANT: Bristol-Myers Squibb Company
 ; TITLE OF INVENTION: AGGRECAN DEGRADING METALLO PROTEASES
 ; FILE REFERENCE: DM6909
 ; CURRENT APPLICATION NUMBER: US/09/122.126B
 ; CURRENT FILING DATE: 1998-07-24
 ; NUMBER OF SEQ ID NOS: 21
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 15
 ; LENGTH: 930
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-122-126B-15

Query Match 16.8%; Score 878; DB 4; Length 930;
 Best Local Similarity 28.9%; Pred. No. 138-69;
 Matches 264; Conservative 116; Mismatches 342; Indels 192; Gaps 39;

QY 135 EYHRGDYVSHIMHRRARRAVSVESL-----HLRLKPRHDMRLTSSSL 186
 Db 50 EYERARPPCHPHPLAGRRSKLVQIDQLYSGCGKVGLYYAGRR-FLDLERDGSV 108
 QY 187 VAPETIVQTLGKTGVQTLPP--EDCFYQGSLSHRNSPSHGKFCBGSTRL--- 240
 Db 109 GTAGFVAGG-----TSAPRHRSHCFYRGV-----DASPRSLAF 146
 QY 241 KLCNSOKCPDSDPRAACAEHNSR-RPFG--RHYKMKPYTOVEADLCKLYCIAGFD 296
 Db 147 DLGGLD-----GFFAVKHRYTLKLLGPMABEEKRGVYDGSARILHY-TRGFS 199
 QY 297 FFFSLSKUKOCPTCS-----EDSRNVCIDGICELSVSTSAHM-POPKED 342
 Db 200 -FEALPFRASCETPASTPEAHEHAPAHNSPGRALASQLDQSNALPAGSGGQ----- 253
 QY 343 LFILPDEYKSLRHKRSLRSHRNEELNVLTVVDDKRMQMHENITTVLTLINMVS 402
 Db 254 -----TWRRRRRSISRARQ-----VELLVADASNARLYG-RGLQHYLLTLASIAN 299
 QY 403 ALF-----KD-----GLMGKDGTRH- 417
 Db 300 RLYSHASIEHILRAVVKVVLGDKSLSVSKNAATLKNPKWQHQNQLGDDBEHY 359
 QY 418 DHAILLTGDIQSWNEBPCDTLGPATSGMCKSVRSCTINEDTGLGIAFTIAHESGHFTG 477
 Db 360 DAATLFTREDLCG--HHSCOTLGADVGTCTSPERSCAVIEDDGLHAFTVAHRIHLLG 417
 QY 478 MTHGBGNMKKSGN-----INSPFTLAGRVFWSNPSRQYLHKFLSTQAICLADQP 532
 Db 418 LSHD--DSKFCETFTGSDTKRLMSIILTSIDASKPNSKTSATITEFLDDGHGNCLLDLP 476
 QY 533 KPVKEYKYPEKLPGELYDANTQKWFGEKAKLCMLDFKIDICKALMCH--RIGRK-CET 589
 Db 477 R--KQILGPEELPGQYDQTCNLNLTGPEYVSCP--GMDVCARLMCAVVRQGMVCLT 531
 QY 590 KFWPAAGTTCGHDMCRGCGQCKYVGDEPK-----PTHGMSDWSSWSFCSRTCCGGVS 644
 Db 532 KKLPAVETGTPCGKGRICLOKGV---DKTKKKYVSTSSHGMSWGSWGSGQCSRCGGVQ 588
 QY 645 HRSRLCTNPKPSHGKTCESGTRTLKLSNQKCPDSDVDFRAAQCAEHNSRFRGRHYK- 703

Search completed: October 28, 2003, 23:47:08
 Job time : 58 secs

Db 589 FAYRHNPNAPRNGRYCTKRAIYRSCSLMPCPNKSGSPRHEQCEAKNGYQSDAKGKWT 648
 QY 704 ---WKP-YTOVEDOPLCKLYCIAEGDFPFPSLSNKVKXGDTPCSEDSNRNVCIDGICERVGC 759
 Db 649 FUEWPKYAGVLPADVCKLCTRAKGTGYVVFSPKAVTDGTECRPYNSVCVRGKCVRTGC 708
 QY 760 DNVLGSDAIVEDGCVGNGNNSACTTHRGLYTKHHHTQYVYHMTIPSGARSIRIYEMNVS 819
 Db 709 DGIIGSKLOYDKGCGGDNSSCTKIVGTENK--KSKGYTDVVRIP#GATHIKVRQFKAK 766
 QY 820 -----TSYISVRNALRYLYNGHWTVDWPGR-YKPSGTTFDYRYSYNEPENIATG--PT 871
 Db 767 DQTRFAYLALKKKKGEYLINGKYMISTSETIIDINGTVNYSWMSHRDQFLHGMGYSAT 826
 QY 872 NETLIVELLFOQNP---GVAMEYSNPLRGTEK-----QPPAQPSYV---W 911
 Db 827 KEILLIVQIL--ATDPTKPLDVRYSFFVPKASTPKVNSVTSGHSNKVGSHTSQFQWVTGPW 884
 QY 912 AIVRSECSVSCGG 925
 Db 885 L-----ACSRCTDTG 894

Db 59 -----EYDLVSAYEVDRHCDYVSHEIMHQRRAVAVSEVESLHLRLKSGSRHDFHVDL 112
 QY 161 RTSSSLVAFGFIQTGLTKGTSVQTLPPEDFCFYQGSLSRSHRNSPSHGKFCGSGSTR 240
 Db 113 RTSSSLVAFGFIQTGLTKGTSVQTLPPEDFCFYQGSLSRSHRNS-SVALSTCGLSGMI 171
 QY 241 KLCNSOKCRDSDVFRAAQCAEHSRRFRG-----HYKWKPYTOVEA-DLCKLYCI 291
 Db 172 R-----TEADYFLRPLPSHLWSKL-GRAAQSSPSHVLYKRSTEPHAGASEVLVT 222
 QY 292 AEGDFFFS--LSNKKVDGTPCSEDSNRVCIIDGICELSVSVSAHMPQPPKEDLFLPDE 349
 Db 223 SRTMELAHQPLHSSDLRLGLP--QKQHFC-----GRRKKYMPQPPKEDLFLPDE 270
 QY 350 YKSLRLHKLRLSHRNEELNVELTVVDDKQMQNHGHEINITVYVLTILNWSALFKD-- 407
 Db 271 YKSLRLHKLRLSHRNEELNVELTVVDDKQMQNHGHEINITVYVLTILNWSALFKDGT 330
 QY 408 -----GLMKGKQDTRHDAILLTGLD 427
 Db 331 IGGNINIAIVGLILLEDQGLVISHADHTLSSFCQWQGLMKGKQDTRHDAILLTGLD 390
 QY 428 ICSWKNPECDTLGFAPISGMCKSYRSCITINEDTGLGLAFTIAHSGHNFMIHDEGNMC 487
 Db 391 ICSWKNPECDTLGFAPISGMCKSYRSCITINEDTGLGLAFTIAHSGHNFMIHDEGNMC 450
 QY 488 KKEGNIIMPTLAGNNGVFSWSPCSQRYLHKFLSTAQAICLADQPKVKEYIKPELPG 547
 Db 451 KKEGNIIMPTLAGNNGVFSWSPCSQRYLHKFLSTAQAICLADQPKVKEYIKPELPG 510
 QY 548 LYDANTCKWQGEKAKLQMLDPFKDKICKALMCHRIKGRKCTKMPAAEGTICGHDMWR 607
 Db 511 LYDANTCKWQGEKAKLQMLDPFKDKICKALMCHRIKGRKCTKMPAAEGTICGHDMWR 570
 QY 608 GGQCVKYGDGPKPTHGNSWSSPCSTCGGVSHRSLCTNPKPSHGKFCGSGSTR 667
 Db 571 GGQCVKYGDGPKPTHGNSWSSPCSTCGGVSHRSLCTNPKPSHGKFCGSGSTR 630
 QY 668 TLKLCNSOKCRDSDVFRAAQCAEHSRRFRGHHYKWKPYTOVEODLCKLYCIAEGDF 727
 Db 631 TLKLCNSOKCRDSDVFRAAQCAEHSRRFRGHHYKWKPYTOVEODLCKLYCIAEGDF 690
 QY 728 FFSLSNKKVDGTPCSEDSNRVCIIDGICERVCDDNVLGSDADEVCGVCNNSACTIHRG 787
 Db 691 FFSLSNKKVDGTPCSEDSNRVCIIDGICERVCDDNVLGSDADEVCGVCNNSACTIHRG 750
 QY 750 LYTKHHNTQYHNTVTPSGARSIRIYEMNVSTYSIVRNALRYVLYNGHWTVDMPGRYK 847
 Db 751 LYTKHHNTQYHNTVTPSGARSIRIYEMNVSTYSIVRNALRYVLYNGHWTVDMPGRYK 810
 QY 848 FSGTTFDYRSYNEPENLIATGPTNETLIVELLFQGRNPGVAWEYSMPRLGTEKOPPAOP 907
 Db 811 FSGTTFDYRSYNEPENLIATGPTNETLIVELLFQGRNPGVAWEYSMPRLGTEKOPPAOP 870
 QY 908 SYTNALVRSECVSCGGGR 926
 Db 871 SYTNALVRSECVSCGGGQ 889

RESULT 2

US-10-296-616-2

; Sequence 2, Application US/10296616
 ; Publication No. US20030129658A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Yamanouchi Pharmaceutical Co., Ltd.
 ; TITLE OF INVENTION: No. US20030129658A1e1 protease
 ; FILE REFERENCE: Y0132PCT-664
 ; CURRENT APPLICATION NUMBER: US/10/296.616
 ; CURRENT FILING DATE: 2002-11-26
 ; PRIOR APPLICATION NUMBER: JP 2000-393372
 ; PRIOR FILING DATE: 2000-12-25
 ; NUMBER OF SEQ ID NOS: 26
 ; SEQ ID NO 2

; LENGTH: 1224
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-296-616-2

Query Match 73.5%; Score 3846.5; DB 16; Length 1224;

Best Local Similarity 75.4%; Pred. No. 0;
 Matches 738; Conservative 25; Mismatches 73; Indels 143; Gaps 13;

QY 1 MKPRARGWGLAALMWLLAQAQVSPGRSHQGRNGSGOLEASPRLLSRGPRRLTAMS 60
 Db 1 MKPRARGWGLAALMWLLAQAQVSPGRSHQGRNGSGOLEASPRLLSRGPRRLTAMS 60
 QY 61 PLFSAAGTCVHRGTRGSGAWFERPASSTRGAAGLDGKGRDNDZAGNRHQQTNTGTENQ 120
 Db 32 PAAAA-----PQSPVPRPPPAERPG-----MNEKG-----58
 QY 121 TLHVLTQYDLVSAYEVDRHCDYVSHEIMHQRRAVAVSEVESLHLRLKSGSRHDFHMDL 180
 Db 59 -----EYDLVSAYEVDRHCDYVSHEIMHQRRAVAVSEVESLHLRLKSGSRHDFHMDL 112
 QY 181 RTSSSLVAFGFIQTGLTKGTSVQTLPPEDFCFYQGSLSRSHRNSPSHGKFCGSGSTR 240
 Db 113 RTSSSLVAFGFIQTGLTKGTSVQTLPPEDFCFYQGSLSRSHRNS-SVALSTCGLSGMI 171
 QY 241 KLCNSOKCRDSDVFRAAQCAEHSRRFRG-----HYKWKPYTOVEA-DLCKLYCI 291
 Db 172 R-----TEADYFLRPLPSHLWSKL-GRAAQSSPSHVLYKRSTEPHAGASEVLVT 222
 QY 292 AEGDFFFS--LSNKKVDGTPCSEDSNRVCIIDGICELSVSVSAHMPQPPKEDLFLPDE 349
 Db 223 SRTMELAHQPLHSSDLRLGLP--QKQHFC-----GRRKKYMPQPPKEDLFLPDE 270
 QY 350 YKSLRLHKLRLSHRNEELNVELTVVDDKQMQNHGHEINITVYVLTILNWSALFKD-- 407
 Db 271 YKSLRLHKLRLSHRNEELNVELTVVDDKQMQNHGHEINITVYVLTILNWSALFKDGT 330
 QY 408 -----GLMKGKQDTRHDAILLTGLD 427
 Db 331 IGGNINIAIVGLILLEDQGLVISHADHTLSSFCQWQGLMKGKQDTRHDAILLTGLD 390
 QY 428 ICSWKNPECDTLGFAPISGMCKSYRSCITINEDTGLGLAFTIAHSGHNFMIHDEGNMC 487
 Db 391 ICSWKNPECDTLGFAPISGMCKSYRSCITINEDTGLGLAFTIAHSGHNFMIHDEGNMC 450
 QY 488 KKEGNIIMPTLAGNNGVFSWSPCSQRYLHKFLSTAQAICLADQPKVKEYIKPELPG 547
 Db 451 KKEGNIIMPTLAGNNGVFSWSPCSQRYLHKFLSTAQAICLADQPKVKEYIKPELPG 510
 QY 548 LYDANTCKWQGEKAKLQMLDPFKDKICKALMCHRIKGRKCTKMPAAEGTICGHDMWR 607
 Db 511 LYDANTCKWQGEKAKLQMLDPFKDKICKALMCHRIKGRKCTKMPAAEGTICGHDMWR 570
 QY 608 GGQCVKYGDGPKPTHGNSWSSPCSTCGGVSHRSLCTNPKPSHGKFCGSGSTR 667
 Db 571 GGQCVKYGDGPKPTHGNSWSSPCSTCGGVSHRSLCTNPKPSHGKFCGSGSTR 630
 QY 668 TLKLCNSOKCRDSDVFRAAQCAEHSRRFRGHHYKWKPYTOVEODLCKLYCIAEGDF 727
 Db 631 TLKLCNSOKCRDSDVFRAAQCAEHSRRFRGHHYKWKPYTOVEODLCKLYCIAEGDF 690
 QY 728 FFSLSNKKVDGTPCSEDSNRVCIIDGICERVCDDNVLGSDADEVCGVCNNSACTIHRG 787
 Db 691 FFSLSNKKVDGTPCSEDSNRVCIIDGICERVCDDNVLGSDADEVCGVCNNSACTIHRG 750
 QY 788 LYTKHHNTQYHNTVTPSGARSIRIYEMNVSTYSIVRNALRYVLYNGHWTVDMPGRYK 847
 Db 751 LYTKHHNTQYHNTVTPSGARSIRIYEMNVSTYSIVRNALRYVLYNGHWTVDMPGRYK 810
 QY 848 FSGTTFDYRSYNEPENLIATGPTNETLIVELLFQGRNPGVAWEYSMPRLGTEKOPPAOP 907
 Db 811 FSGTTFDYRSYNEPENLIATGPTNETLIVELLFQGRNPGVAWEYSMPRLGTEKOPPAOP 870

QY 908 SYTWAIVRSCSVSCGGG 926
 |||||
 Db 871 SYTWAIVRSCSVSCGGG 889

RESULT 3

US-10-240-545A-2
 ; Sequence 2, Application US/10240545A
 ; Publication No. US20030185828A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Yamanouchi Pharmaceutical Co., Ltd.
 ; TITLE OF INVENTION: No. US20030185828A1el aggrrecanase
 ; FILE REFERENCE: 08959.0002
 ; CURRENT APPLICATION NUMBER: US/10/240,545A
 ; CURRENT FILING DATE: 2002-10-02
 ; PRIOR APPLICATION NUMBER: PCT/JP01/11033
 ; PRIOR FILING DATE: 2001-12-17
 ; PRIOR APPLICATION NUMBER: JP 2000-384300
 ; PRIOR FILING DATE: 2000-12-18
 ; NUMBER OF SEQ ID NOS: 26
 ; SEQ ID NO 2
 ; LENGTH: 1221
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-240-545A-2

Query Match 46.5%; Score 2435.5; DB 12; Length 1221;
 Best Local Similarity 51.3%; Pred. No. 1.2e-207;
 Matches 466; Conservative 118; Mismatches 195; Indels 129; Gaps 17;
 QY 85 ASSTGGAAGLDGKGRDDEAGNHRSSQQTNTGTENQTLHLVLTQYDLVSAVEVDHGRDYVS 144
 |||||
 Db 47 ASDSSSGAGLN-----DDYVFTPEVDSAGSVIS 77
 QY 145 HEIMHQRRAVAVSEVESLHLRLKGRPHDHWLRTSSSLVAPGFTVQTLGKTGYSV 204
 |||||
 Db 78 HDILHNGRKKES-AQNAASSLHYRPSAFQELHLELK-PSAILSSHFIQVVLCKDGASBT 135
 QY 205 QTLPPDFCFYQGSLSHRNSPHGKFCBEGSTRTTLKLCNSQ-----KCPRDSVDFRAAQC 260
 |||||
 Db 136 QK-PEVQCFYQGFIRND-SSSSVAVSTCAGLSGLIRTKNEFLISPLFQ-----LLA 186
 QY 261 AEHNSRRFRGHYKMKPTQVDEADLCKLYCIAGGDFPFSLSNKVDG-----TPCSBDS 315
 |||||
 Db 187 QEHNYSSPAG-HHPVLYKRTABEKTQRYGPG-----SCRNYGYSPSHIPHASQS 238
 QY 316 R-----NVCIDGICELSVSVSTSAHMPQPKEDFLFIPDEYKSLRHKRSLRS 363
 |||||
 Db 239 RETEYHRRLOKQHF-----GRKKYAPKPTEDTYLRFDEYSSGSRPRASAGS 289
 QY 364 HRNEELNVTLLVVDDKQWQNHENITVYLLAMVTSALFKDG-----KCPRDSVDFRAAQC 260
 |||||
 Db 290 QKG--LNVETLWADKQVKNHKGKGNVTTLVWNNVSGLFKDGDTIGSDINVVVLSIL 347
 QY 409 -----LMGKQGRHDAHLLATGLDICSWKNEPCDTLGF 441
 Db 348 LEQFQGLLINHADQSLNSFCWQSAIIGKNGKRDHAILLATGFDICSWKNEPCDTLGF 407
 QY 442 APISGMCSTYRSTINEDTGLGLAFTIAHSGHNFQMTIHGEGNMCKKSGEINTMSPTIAG 501
 Db 408 APISGMCSTYRSTINEDTGLGLAFTIAHSGHNFQMTIHGEGNFCRAKGEINTMSPTIAG 467
 QY 502 RNVGFSWSPCSRQYVHLKFLSTAAQICLADQPKPKYKPKLPQBELVDANTOCKWQFGE 561
 Db 468 NNGVFSNSCSRQYVHLKFLSTAAQICLADQPKPKYKPKLPQBELVDANTOCKWQFGE 527
 QY 562 KAKLMLQPKDTCALWCHRIQRKCTKFPWPAABETICGHDMKRCQGVQGVGDEGPKP 621
 Db 528 KAKLCSLGFVKDICKSLWCHVGRCTKFPWPAABETICGHDMKRCQGVQGVGDEGPKP 587
 QY 622 THGHWDSSWSPCSTCGGVSHRRICTNPKPSHGKFCBEGSTRTTLKLCNSQCPRDS 681
 |||||
 Db 588 IHQWAMSKWSECSRSTCGGVVQBRCHNNPKPQVGLFCPGSSRIYQLCNINFCNEN 647

RESULT 4

US-10-226-560-2
 ; Sequence 2, Application US/10226560
 ; Publication No. US20030050464A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hu, Yi
 ; APPLICANT: Nepomutichy, Boris
 ; APPLICANT: Pakke, Wade
 ; APPLICANT: Padoe, John
 ; TITLE OF INVENTION: No. US20030050464A1el Human Proteases and Polynucleotides Encod
 ; FILE REFERENCE: 0272-USA
 ; CURRENT APPLICATION NUMBER: US/10/226,560
 ; CURRENT FILING DATE: 2002-08-22
 ; PRIOR APPLICATION NUMBER: US 60/314,049
 ; PRIOR FILING DATE: 2001-08-22
 ; PRIOR APPLICATION NUMBER: US 09/917,614
 ; PRIOR FILING DATE: 2001-07-27
 ; PRIOR APPLICATION NUMBER: US 60/221,644
 ; PRIOR FILING DATE: 2000-07-28
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 862
 ; TYPE: PRT
 ; ORGANISM: homo sapiens
 ; US-10-226-560-2

Query Match 43.3%; Score 2265; DB 15; Length 862;
 Best Local Similarity 50.7%; Pred. No. 1.2e-192;
 Matches 435; Conservative 107; Mismatches 194; Indels 122; Gaps 14;
 QY 85 ASSTGGAAGLDGKGRDDEAGNHRSSQQTNTGTENQTLHLVLTQYDLVSAVEVDHGRDYVS 144
 |||||
 Db 47 ASDSSSGAGLN-----DDYVFTPEVDSAGSVIS 77
 QY 145 HEIMHQRRAVAVSEVESLHLRLKGRPHDHWLRTSSSLVAPGFTVQTLGKTGYSV 204
 |||||
 Db 78 HDILHNGRKKES-AQNAASSLHYRPSAFQELHLELK-PSAILSSHFIQVVLCKDGASBT 135
 QY 205 QTLPPDFCFYQGSLSHRNSPHGKFCBEGSTRTTLKLCNSQ-----KCPRDSVDFRAAQC 260
 |||||
 Db 136 QK-PEVQCFYQGFIRND-SSSSVAVSTCAGLSGLIRTKNEFLISPLFQ-----LLA 186
 QY 261 AEHNSRRFRGHYKMKPTQVDEADLCKLYCIAGGDFPFSLSNKVDG-----TPCSBDS 315
 |||||
 Db 187 QEHNYSSPAG-HHPVLYKRTABEKTQRYGPG-----SCRNYGYSPSHIPHASQS 238
 QY 316 R-----NVCIDGICELSVSVSTSAHMPQPKEDFLFIPDEYKSLRHKRSLRS 363
 |||||
 Db 239 RETEYHRRLOKQHF-----GRKKYAPKPTEDTYLRFDEYSSGSRPRASAGS 289

364	QY	HRNBEUNVETLVVDKMMQKHGHENTTYVLTILNKNVSALEKQO	408
365	QY		
366	QY		
367	QY		
368	QY		
369	QY		
370	QY		
371	QY		
372	QY		
373	QY		
374	QY		
375	QY		
376	QY		
377	QY		
378	QY		
379	QY		
380	QY		
381	QY		
382	QY		
383	QY		
384	QY		
385	QY		
386	QY		
387	QY		
388	QY		
389	QY		
390	QY		
391	QY		
392	QY		
393	QY		
394	QY		
395	QY		
396	QY		
397	QY		
398	QY		
399	QY		
400	QY		
401	QY		
402	QY		
403	QY		
404	QY		
405	QY		
406	QY		
407	QY		
408	QY		
409	QY		
410	QY		
411	QY		
412	QY		
413	QY		
414	QY		
415	QY		
416	QY		
417	QY		
418	QY		
419	QY		
420	QY		
421	QY		
422	QY		
423	QY		
424	QY		
425	QY		
426	QY		
427	QY		
428	QY		
429	QY		
430	QY		
431	QY		
432	QY		
433	QY		
434	QY		
435	QY		
436	QY		
437	QY		
438	QY		
439	QY		
440	QY		
441	QY		
442	QY		
443	QY		
444	QY		
445	QY		
446	QY		
447	QY		
448	QY		
449	QY		
450	QY		
451	QY		
452	QY		
453	QY		
454	QY		
455	QY		
456	QY		
457	QY		
458	QY		
459	QY		
460	QY		
461	QY		
462	QY		
463	QY		
464	QY		
465	QY		
466	QY		
467	QY		
468	QY		
469	QY		
470	QY		
471	QY		
472	QY		
473	QY		
474	QY		
475	QY		
476	QY		
477	QY		
478	QY		
479	QY		
480	QY		
481	QY		
482	QY		
483	QY		
484	QY		
485	QY		
486	QY		
487	QY		
488	QY		
489	QY		
490	QY		

RESULT 5

```

US/10-217-774-2
Sequence 2, Application US/10217774
Publication No. US20020193583A1
GENERAL INFORMATION:
APPLICANT: Fildip, Earl Johan
APPLICANT: Hilburn, Erin
TITLE OF INVENTION: No. US20020193583A1
TITLE OF INVENTION: Same
FILE NUMBER: LEX-0219-US3
CURRENT FILING DATE: 2002-08-12
PRIORITY FILING DATE: 2002-09-12
PRIORITY APPLICATION NUMBER: US/09/331
PRIORITY FILING DATE: 2001-08-14
PRIORITY APPLICATION NUMBER: US 60/221
PRIORITY FILING DATE: 2000-08-16
NUMBER OF SEQ ID NOS: 5
SOFTWARE: Fastseq for Windows Version 1.0
SEQ ID NO 2
LENGTH: 491
ORGANISM: homo sapiens
US/10-217-774-2

```

```

Query Match      29.5%; Score 1543.5; DB 14; Length 491;
Best Local Similarity 57.7%; Pred. No. 1.4e-128;
Matches 333; Conservative 26; Mismatches 75; Indels 143; Gaps 13;

QY 1 MKPRARGMGALAAWMLLAQVAEQVSPGRSHQRNGRSGQLEASPPRLSLRGPRRITAMS 60
    |||
DB 1 MKPRARGMGALAAWMLLAQVAEQAP-----ACNMG 31

```

[illegible]

RESIST 6

```

US09-780-043C-1
1 Sequence 1, Application US/0978043C
2 Patent No. US20020107361A1
3 GENERAL INFORMATION:
4 APPLICANT: Heller, Renu
5 APPLICANT: Zuo, Fengrong
6 ATTORNEY: Kowowski, PUS20020107361A1
7 TITLE OF INVENTION: Thrombospondin Domain
8 TITLE OF INVENTION: Encoding the Same
9 FILE REFERENCE: ROCH-004
10 CURRENT APPLICATION NUMBER: US/09/780, 043C
11 PRIOR FILING DATE: 2001-02-16
12 PRIOR APPLICATION NUMBER: 60/184, 152
13 PRIOR FILING DATE: 2000-02-18
14 NUMBER OF SEQ ID NOS: 10
15 SOFTWARE: FastSeq for Windows Version 4.0
16 SEQ ID NO 1
17 LENGTH: 959
18 TYPE: PPT
19 ORGANISM: human
20 FEATURE:
21 NAME/KEY: VARIANT
22 LOCATION: 909..921..
23 OTHER INFORMATION: Xaa = Any Amino Acid
24 US-09-780-043C-1

```

```

Query Match      26.6%; Score 1391; DB 10; Length 959;
Best Local Similarity 34.2%; Pred. NO. 1.2e-114;
Matches 308; Conservative 141; Mismatches 309; Indels 142; Gaps 24;

QY 108 HRSQOQNTGNTNTHLVLTQDVLVSAYEDVHGDPYVSHIWH--HQRRRRAV-----AVSE 161
                                     |||
Db 22 HSDHRLSVSSQREFLVLEHYQTPIRVNDQAGPISFTVNDKSHRRSRMSDPIPOCA 81
                                     |||

```

```

QY 162 VESLHLKGRPHRDHMDLTSSIVAPGFIYOTLGTGTGKTSVOTLPPE-----DFCFY 215
DB 82 VSKLFFKLSAYGKHFLNLTLNTDFVSKHFTVEYMGKG-----PQKHDFLNDCHY 133
QY 216 QGSLSHRNPSHGKFCESGTRTLKLCNSQKCPDRSDVF-----RAAOCAENRRFRG 270
DB 134 TGYLDQRSTTKVALSNVGLHGV-----ATEDEEYFIEPLKNITTEDSKHFSYENGH 186
QY 271 RHVKKPYTQVEADLC-KLYCIAEGDFPFSLSNKVDGTP-CSEDSRNVCIDIGICELSV 328
DB 187 PHVIVKSAQQORHLHYDHSCH-----GVSDFTSRGKPMWLNDTSTVSY-----SLPI 233
QY 329 VSTSAMHPQPKEDFILPDEYKSLRHSLLSRHNEELAVETLVVVDKMMQNHGHE 388
DB 234 NNTIHH-----ROKRSV-----SIERFVELVADGMMVGYGRK 269
QY 389 NITTVILTNMVSALFKDGLMG-----411
DB 270 DIEHILSVNVIIVAKLYRDSLSGNVNIIVARLIVLTEDQPNLEINHHAOKSLDSFCKWQ 329
QY 412 -----KDG-----TRDHAILLGLDLCMSKNPCDITLGPAPISGMSKYSRCTI 456
DB 330 KSLSHQSDGNTIPENGIAHHDNAVLITRYDICTYKNKPCGTGLASVAGMCEPERSCSI 389
QY 457 NEDTGLGAPTTIAHSGHNTFMHDEGNMC--KKSE--GNINSPTLAGRNGVFSKPCSR 513
DB 390 NEDIGLSAFTIAEIGHNFGMHDGICSGTGKGEAAKMAAHTANTNPFWSACS 449
QY 514 QYLRKFLSTQAICLADOPKPVKYEKPEKLPGLYDANTQCKWQGEKAKMLDPKKD 573
DB 450 DYTSLFSGRGTCLDNEP-PKRDFLYPAVAPQVVDADQCFRQFYGATSRQCKYV--E 505
QY 574 ICALMCHIRGRKCTKMPAAEGTICG-----HDMWCRGGQCVKYVDEGPKPTHGWSDW 629
DB 506 VCEBLCLSKSNRCVNTNIPAAEGTLCTQGNIEKMGWCYQDGVPRG-TWQPSIDGNGPW 564
QY 630 SSMSPCSTCGGVSRRSELCTNPKPSPGKFCGEGSTTLKLCNSQKCPDRSDVFRAQC 689
DB 565 SLMGECSTCGGVSRRSELCHDPSAPSGGKCYCLGERKRYSCNTDPCPLGSDPREKQC 624
QY 690 AEHNSRRFRGRHYKWKPYTOVEDQDLCKLYCIAEGDFPFSLSNKVDGTPCSEDSNRVC 749
DB 625 ADFDNNPFRGKYNNKPYTGGVKP-CALNCLAEGVNFTYTERAPAVIDGTQCNADSLDIC 683
QY 750 IDGICERVGCDNLGSDADEVCGVGNNGNSACTTHRGLYTKHHTNYYHMTIPSGAR 809
DB 684 INGECKHVGCDNLGSDAREDCRVCGDGGSTCDIEGFNDLSFRGGYMEVVQIPRGSV 743
QY 810 SIRIEMNVSTYSIVRNALRYLNGHWTVDWPGRYKFGSTTFDYRRSYNEPENLIATG 869
DB 744 HIEVREAVMSKVIALKSEGDDYINGAWITDPRKFDVAGTAFHYKPTDEPESLEALG 803
QY 870 PTNETLIVELLFQGRNPGVAMEYSMP--RLGTEKOPPAQPSYTWAIVR--SECSVSCGGR 926
DB 804 PTSENLIWVLLQEQNLGIRYKFNVEITRTGSDN---EVGFTMNHQPSWSECSATCAGK 860

```

RESULT 7

```

US-09-963-791-2
; Sequence 2, Application US/09963791
; Patent No. US2002012013A1
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedlich, Glenn
; APPLICANT: Scoville, John
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. US2002012013A1a1 Human Proteases and Polynucleotides Encodir
; FILE REFERENCE: LEX-0108-USA
; CURRENT APPLICATION NUMBER: US/09/963,791
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: US 60/169,769

```

```

; PRIOR FILING DATE: 1999-12-09
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 908
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-963-791-2

```

```

Query Match 26.5%; Score 1390; DB 10; Length 908;
Best Local Similarity 34.2%; Pred. No. 1.4e-114;
Matches 308; Conservative 141; Mismatches 309; Indels 142; Gaps 24;

```

```

QY 108 HRSQQTNTGENTLHVLTQYDLVSAYEVDHGRDYVSHIMH--HQRRAV-----AVSE 161
DB 22 HSDHLLSYSSQBEFTLYLEHYQTIPRVQNGAFLSTVYNDKHSRRSRMDPDPQQA 81
QY 162 VESLHLKGRPHRDHMDLTSSIVAPGFIYOTLGTGTGKTSVOTLPPE-----DFCFY 215
DB 82 VSKLFFKLSAYGKHFLNLTLNTDFVSKHFTVEYMGKG-----PQKHDFLNDCHY 133
QY 216 QGSLSHRNPSHGKFCESGTRTLKLCNSQKCPDRSDVF-----RAAOCAENRRFRG 270
DB 134 TGYLDQRSTTKVALSNVGLHGV-----ATEDEEYFIEPLKNITTEDSKHFSYENGH 186
QY 271 RHVKKPYTQVEADLC-KLYCIAEGDFPFSLSNKVDGTP-CSEDSRNVCIDIGICELSV 328
DB 187 PHVIVKSAQQORHLHYDHSCH-----GVSDFTSRGKPMWLNDTSTVSY-----SLPI 233
QY 329 VSTSAMHPQPKEDFILPDEYKSLRHSLLSRHNEELAVETLVVVDKMMQNHGHE 388
DB 234 NNTIHH-----ROKRSV-----SIERFVELVADGMMVGYGRK 269
QY 389 NITTVILTNMVSALFKDGLMG-----411
DB 270 DIEHILSVNVIIVAKLYRDSLSGNVNIIVARLIVLTEDQPNLEINHHAOKSLDSFCKWQ 329
QY 412 -----KDG-----TRDHAILLGLDLCMSKNPCDITLGPAPISGMSKYSRCTI 456
DB 330 KSLSHQSDGNTIPENGIAHHDNAVLITRYDICTYKNKPCGTGLASVAGMCEPERSCSI 389
QY 457 NEDTGLGAPTTIAHSGHNTFMHDEGNMC--KKSE--GNINSPTLAGRNGVFSKPCSR 513
DB 390 NEDIGLSAFTIAEIGHNFGMHDGICSGTGKGEAAKMAAHTANTNPFWSACS 449
QY 514 QYLRKFLSTQAICLADOPKPVKYEKPEKLPGLYDANTQCKWQGEKAKMLDPKKD 573
DB 450 DYTSLFSGRGTCLDNEP-PKRDFLYPAVAPQVVDADQCFRQFYGATSRQCKYV--E 505
QY 574 ICALMCHIRGRKCTKMPAAEGTICG-----HDMWCRGGQCVKYVDEGPKPTHGWSDW 629
DB 506 VCEBLCLSKSNRCVNTNIPAAEGTLCTQGNIEKMGWCYQDGVPRG-TWQPSIDGNGPW 564
QY 630 SSMSPCSTCGGVSRRSELCTNPKPSPGKFCGEGSTTLKLCNSQKCPDRSDVFRAQC 689
DB 565 SLMGECSTCGGVSRRSELCHDPSAPSGGKCYCLGERKRYSCNTDPCPLGSDPREKQC 624
QY 690 AEHNSRRFRGRHYKWKPYTOVEDQDLCKLYCIAEGDFPFSLSNKVDGTPCSEDSNRVC 749
DB 625 ADFDNNPFRGKYNNKPYTGGVKP-CALNCLAEGVNFTYTERAPAVIDGTQCNADSLDIC 683
QY 750 IDGICERVGCDNLGSDADEVCGVGNNGNSACTTHRGLYTKHHTNYYHMTIPSGAR 809
DB 684 INGECKHVGCDNLGSDAREDCRVCGDGGSTCDIEGFNDLSFRGGYMEVVQIPRGSV 743
QY 810 SIRIEMNVSTYSIVRNALRYLNGHWTVDWPGRYKFGSTTFDYRRSYNEPENLIATG 869
DB 744 HIEVREAVMSKVIALKSEGDDYINGAWITDPRKFDVAGTAFHYKPTDEPESLEALG 803
QY 870 PTNETLIVELLFQGRNPGVAMEYSMP--RLGTEKOPPAQPSYTWAIVR--SECSVSCGGR 926
DB 804 PTSENLIWVLLQEQNLGIRYKFNVEITRTGSDN---EVGFTMNHQPSWSECSATCAGK 860

```

```

RESULT 8
US-10-103-377C-2
; Sequence 2, Application US/10103377C
; Publication No. US20030073098A1
; GENERAL INFORMATION:
; APPLICANT: Bandaru, Rajasekhar
; TITLE OF INVENTION: 65577, A Human Matrix
; FILE REFERENCE: Metalloproteinase and Uses Thereof
; CURRENT APPLICATION NUMBER: US/10/103,377C
; PRIOR FILING DATE: 2002-08-26
; PRIOR FILING DATE: 2001-03-23
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 684
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-103-377C-2

Query Match      26.3%; Score 1376; DB 15; Length 684;
Best Local Similarity 35.8%; Pred. No. 1.6e-113;
Matches 303; Conservative 93; Mismatches 154; Indels 316; Gaps 14;

QY 85 ASSTGAGAGLGGKGMDEAGNHRSSQQTNTGTENLHVLVQYDLVSAIEVDHRGDYVS 144
DQ 47 ASDSSGASGLN-----DDXVFVTPEVDGAGSYIS 77
QY 145 HEIMHORRRRAVAVSEVSLHLKGRPHDPMDLTSSSLVAPGFIQTLGKTVKSV 204
DQ 78 HDILNGRKRGS-AQNAKSLHYPSAFQOELHLLELK-PSATLSHFYIVVLGKGASLT 135
QY 205 QTLPPEDFQYGGSLRHNRNPSGSGKPCGSGTTLKLNCSKCPKPSVDVFAAQ----- 259
DQ 136 QK-FEVQCCFYQGFIRND-SSSVAVATCAGSLGIRTKNEFLISLPQLLAQHNRSS 193
QY 260 -----CAHNSRPR-----GRYK-WKPYQVTEADLCKLYCIAGDFPFPSL 301
DQ 194 PAGHPHVLKTAETKRYGIFGSGRVRPGYSPSHIPHA-----SQ 237
QY 302 SNVKY-----DGTPTCEDSRNVGIDGELSVSTSNKMPKPKDLFILDEYKSLR 355
DQ 238 SRETEYHRRLLKSLIFVDARNV-----YAPKPTEDTLRFDYGGSR 282
QY 356 HYSLLSHSRNEELNVETLVVVKQKMQHGHENITTYVLTLNMYSLAFKDG----- 408
DQ 283 PRSAGSKQG--LNVETLVVADKWKVYKKGKGNVTYTLTVMNWSGLFKDGTIGSDIN 343
QY 409 -----LWGDQGRDHDHAILLTGLDICSWN 430
DQ 341 VVVVSLILLAEQEPGLLIINHADQSILNSFCQWQALIGRNGKRDHAILLTGFDICSWN 400
QY 434 EPCDTLGFAPISGMSKYRSCITNEDTGLGLAFTIAHSGHNFQMIHDEGNMCKSGEN 493
DQ 401 EPCDTLGFAPISGMSKYRSCITNEDTGLGLAFTIAHSGH----- 441
QY 494 IMSPTLAGRNGVFSMPCSRQYLHKFLSTAQAICLAQPKPKYKPKYKPKLPGELYDANT 553
DQ 442 ----- 441
QY 554 QCKWQFGEKAKLMLDFKDIKALWCHIRGKCKETKMPAAEGTICGHDMWCRGQCVK 613
DQ 442 ----- 441
QY 614 YGDEQPKPTHGWSMWSMSPCSRTCCGGVSHSRSLCTNPKPSHGKFGCEGSTRKLKCN 673
DQ 442 -----KPYGGLFPCGSSRIYOLCN 461
QY 674 SQKCPKPSVDFAAQCAHNSRRFRGRHYKWKPYQVEDQDLCKLYCIAGDFPFPSLSN 733
DQ 462 INPCNENSLDPAQAQCAETNSKPFGRFGYQWKPYTKVEEEDRCKLYCAENFEFFPANSJ 521

```

```

QY 734 KYKDGTPCSDSRNVCTDGTICERVCCDNVLGSDAVEDVCGVCGNCGNSACTIHRGLYTKKH 793
DQ 522 KYKDGTPCSDSRNVCTDGTICERVCCDNVLGSDAVEDVCGVCGNCGNSACTIHRGLYTKKH 581
QY 794 HTNOYTHMVTIPSGARSIRIYEMNVSTSYISVNRALRRYVLYNGHWTVDMGRKFSJTFP 853
DQ 582 KANEYTPVLLIPAGARSIEIQELQVSSSYLAVERSLSKQYLYLGGWSIDWPGFEFFACTIF 641
QY 854 DYRRSYNEPENLIATGPTNLTILVEL 879
DQ 642 EYORSNRPRLYAPGPTNETILVFEV 667

RESULT 9
US-09-963-791-24
; Sequence 24, Application US/09963791
; Patent No. US20020120113A1
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Scoville, John
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. US20020120113A1el Human Proteases and Polynucleotides Encod
; FILE REFERENCE: LEX-0105-USA
; CURRENT APPLICATION NUMBER: US/09/963,791
; PRIOR FILING DATE: 2000-12-08
; PRIOR FILING DATE: 1999-12-09
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 757
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-963-791-24

Query Match      25.2%; Score 1319; DB 10; Length 757;
Best Local Similarity 40.5%; Pred. No. 2.3e-108;
Matches 257; Conservative 106; Mismatches 195; Indels 76; Gaps 13;

```

```

QY 355 RHRSLLSRHNEELNVETLVVVKQKMQHGHENITTYVLTLNMYSLAFKDLGMLG--- 411
DQ 90 RQKRSV-----SIRFVELVADKWKVYKKGKDI EHLISVMKIVAKLIRLSLGRVV 144
QY 412 -----KDG-----TBDHAIL 422
DQ 145 NIIVARLVLTEQPLEINLHNAKSLDSFCWQKSLSHQSDGNTIPENGTAHIDNVL 204
QY 423 LTGLDICSWNPCDTLGFAPISGMSKYRSCITNEDTGLGLAFTIAHSGHNFQMIHDEGN 482
DQ 205 IIRYDICTYKPKCGTGLASVAGNCPERSCSINEDIGLGSAPFLIAHIGNFGNHDG 264
QY 483 EGNWC--KKSE-GNIMSPTLAGRNGVFSMPCSRQYLHKFLSTAQAICLAQPKPKYK 539
DQ 265 IGNSCGTKGHEAKLAAAHITANTNPFMSACSADYIITSFLDSGRGTCLDNEP-PKDFL 323
QY 540 YTEKLPGLYDANTCKWQFGEKAKLMLDFKDIKALWCHIRGKCKETKMPAAEGTI 599
DQ 324 YFAVAPQVYDADEQCFQVGTATSRQCKYK--EVCRELWCLSKSNRCVTSNIPAAAGTL 380
QY 600 CG---HDMWCRGQCVKYGDEGPKPTHGWSMWSMSPCSRTCCGGVSHSRSLCTNPKP 655
DQ 381 CQTGNIEKWCYQGDVPEFG-TWQSIDGNGWGPMSLWGCSCRTCCGGVSSSLRHCDSPAP 439
QY 656 SHHGKPCBSGSTRKLKCNQKCPKPSVDFAAQCAHNSRRFRGRHYKWKPYQVEDQDL 715
DQ 440 SGGKYLCLGRKRYKRSNTDPCPLGSRDFREKQCAFDNNMPKRYKYNKPYTGGVKK- 498
QY 716 CKLYCIAEGDFPFPSLSNCKVKGDTPCSDSRNVCTDGTICERVCCDNVLGSDAVEDVCGVC 775

```

439	DB	CALNCLAEAGSYNFTTERAPAVIDGTQCNAOSLDICINGCEKHVGCNLTGSDAREDCRVC	558
776	QY	NNNSACTIHRGLVTRKHHHTNQYVHWVTPGASRIRIVEMVSTVSTVSRNLRRYLNL	835
559	DB	GGDSTCDIAIGFNDLSLPGGYMEVWVQIPGSHVIEVRAVMKNYIAIKSEGDDYYIN	610
836	QY	GHWTVMPGRKVKFGSTTFOYRSRYNPENLITGPTNETLIVELLFGRRNPGVAWYXNP	895
619	DB	GAMTIDWPKEVDAGTAFAFYKHPDPEFSELAGETSENLIWVLLQELGIRYKFNVP	678
896	QY	--RLGTKEQFPAQPSVTYIAVR--SECSVSCGGR	926
679	DB	ITRTGSGDNL--EYGTWNHPWPSECSATCAGK	709

```

RESULT 10
US-09-788-043C-5
/ SEQUENCE 5, Application US/09788043C
/ Patent No. US20020107361A1
/ GENERAL INFORMATION:
/ APPLICANT: Heller, Renu
/ APPLICANT: Zuo, Fengrong
/ APPLICANT: Klonowski, Paul
/ TITLE OF INVENTION: No. US20020107361A1el Metalloproteases Having
/ TITLE OF INVENTION: Thrombospondin Domains and Nucleic Acid Compositions
/ TITLE OF INVENTION: Encoding the Same
/ FILE REFERENCES: ROCH-004
/ CURRENT APPLICATION NUMBER: US/09/788,043C
/ CURRENT FILING DATE: 2001-03-16
/ PRIOR APPLICATION NUMBER: 60/184,152
/ PRIOR FILING DATE: 2000-02-18
/ NUMBER OF SEQ ID NOS: 0
/ SOFTWARE: SeqMan 5 FastSeq for Windows Version 4.0
/ SEQ ID NO: 5
/ LENGTH: 1690
/ TYPE: prt
/ ORGANISM: human
US-09-788-043C-5

```

```

417 FMSPOLLYDAAPLWTSRCSQYITRFLDRGWGLCLDDPPAKOIID--FPSVPGVLVDY 477
552 NTOCKWOFBKAKLWOLFDDKICKALWCHRIGKCEKTFKPAEAGTICGHUWMCRCGQC 611
475 SHQCRLOQVAYSAC--EDMONVCHTLAC--SVGTTCBSKLDAAVDGTRCGENKMKCLSGEC 531
612 VKYGDGPKPTHGHWDSSWSPSCSTCGGVGSHRSRLCTNPKPSHGGAFCGSGTRTKL 671
532 VPVGFPR--PEAVDGSNGSANSICSRSCGMGVQSAERQCTOPTPKYKRGVYCVGERKFRLL 590
672 CNSQKCFRSDVDFRAAQCAEHNSRFRGRHYKWKPTQVQEDQDLCKLYCIAEGDFDFESL 731
591 CNLQACFAGRESFRHVQCSHFDAMLYKGRLLHTVPV--VNDVNPCLHCRPANEFYAEKL 648
732 SNKVKDGTGPCE--DSRNVCTDGI6RSGVCONVLGSDAVEDVCVCGNNSACTIHRGLY 789
649 RDAVWDGTCFQVRA5RDLCTNGICKNVQCOFEDSGAMEDRCGVCHNGSGTCHTVSGTF 708
790 TKHHYNYOYHNVTIPSGARSIRYENNTSYISVSRNA--LRYYLNGHWTVDMPGRYKF 748
709 EBAEGLG-YVDVGLIPAGARBIRIQEVA5AANFLAR5EDPEKYFLNGOWTIQMGDYQV 767
849 SGTTFDYRRSYNEPENLIATGFTNETLI1VELLFQGRNPGVAMEYSMPRL--GTEKQPPAQ 906
768 AGTFTFYARRGNW--ENLTSPOGTKEPVIQLLFOESNPGVHYVXTHREAGGHDEVP-- 824
907 PSYTNAL-VRSEKSVSCGGG 925
825 PVFSWHYGMWTKTCTVTCGRG 844

RESULT 11
US-09-858-081-9
/ Sequence 9, Application US/09858081
/ Patent NO. US20020072490A1
/ GENERAL INFORMATION:
/ APPLICANT: Kapeller-Libermann, Rosana
/ APPLICANT: Cook, William James
/ APPLICANT: Silos-Santiago, Immaculada
/ TITLE OF INVENTION: 33428, A NOVEL HUMAN METALLOPROTEASE
/ TITLE OF INVENTION: FAMILY MEMBER AND USES THEREOF
/ FILE REFERENCE: 10448-049001
/ CURRENT APPLICATION NUMBER: US/09/858,081
/ PRIORITY FILING DATE: 2001-05-15
/ PRIOR APPLICATION NUMBER: 60/204,159
/ PRIORITY FILING DATE: 2000-05-15
/ PRIOR APPLICATION NUMBER: 60/204,160
/ PRIORITY FILING DATE: 2000-05-15
/ NUMBER OF SEQ. ID NOS: 14
/ SOFTWARE: Fast-Seq for Windows Version 4.0
/ SEQ ID NO 9
/ LENGTH: 1044
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-858-081-9

```

Query Match	23.7%	Score 1243;	DB 9;	Length 1044;
Best Local Similarity	31.4%;	Prod.No. 2.2e-101;		
Matches	295;	Conservative 133;	Mismatches 317;	Indels 194;
Gaps	24;			
QY	84	PASSPTGACGLOGKRDMDAGNHRSSQQTGTENQTLVLVTOYDLSVAVEVDHRRGVV	143	
DB	3	PAQILFWALAL--GLGLMFEVTHAFRSQ-----DEFLSSLEYSIEIAPPRVDHNGALL	54	
QY	144	SHEIMHQRRRAVAVSVESLSHLRLKGRPHDFMDLRTSSLSVAFCFIVQLTKGTGKS	203	
DB	55	AFSPPPPRQRGTGTAESLSFYKVASPTSHLLNLRSSRLLAGHSVEYMTREGIAM	114	
QY	204	VQTLPPDPFCYQGLSLRSHRNS-----PSHGKGFCEGSTR	238	
DB	115	QRAAREPH--CLYAGHLQQQASSSSHVAISTCCGLLHGLIVADEEYLIEPLRGGP--	168	
QY	239	TLKLNCSOKCPKDS--VDFAAAQ-----CAEHSNRFRGRYKWKYQVQEAADLC	286	

Db 169 -----RSPESGPHVYKSSLRHPLDTCAGVRDEKPKGRPWMLRTLKPPAR-- 218
 Qy 287 KLYCIABGDFPFSSLSNKKVKGDTGTPCSDSRNNVDCIGELSVSVSAHMPQPPKEDLFL 346
 Db 219 -----PLGNTERGQP-----GL----- 231
 Qy 347 PDEYKSLRHSLSLRHSHREELNVELTVVQKQWQNHGHENITTVYTLNMMVSALEK 406
 Db 232 -----KSVSR-----ERYVETLVADKMWVAYHGRDRVQYVLAIMNIAKLQ 276
 Qy 407 DGLMG-----KDG- 414
 Db 277 DSSLGTVNLVTRLLITLTDQPTLEITHHAGSLDSCFKWQKSVNHSHGNAIPENG 336
 Qy 415 TRHDHAILTGIDICSWNKEPCDTLGFAPISGKSVRSCTINEDTGLGTAFTIAHESGH 474
 Db 337 ANHDTAVLITRYDICIYKPKCTGLAPVGGCERSCSVNEDIGLATAFTIAHETGH 396
 Qy 475 NFGMTHDGEKMC---KKSEGNIMSPTLAGRNQVFSMSPCSQYLRKFLSTQAICLADQ 531
 Db 397 TFGMNDHGVNCGARGQDPFAMAAHITMKTNPFWSSCSRDYITSLDGLGLCLNRR 456
 Qy 532 PRPVEKYEPKLPGLDYDANTQCKWQFGEKAKLCLMDFKDICKALWCHRIGRCKETRF 591
 Db 457 P-PRODVPYTPAQYADABQCFQGVKRSCKYK---EVCSELWCLSKNRCITNS 512
 Qy 592 MPAAGETIC-GHDM---WCRGGQCVKYVGDGPKPTHGHWSMSSWSPCSTCGGVSHRS 647
 Db 648 RLCTNPKPSHGKFCESGTRTLKNSOKCPROSDVDFRAAQCAEHNSRRFRGRHYKWKPY 707
 Qy 572 RHCDSPRTTGGYKCLGRHRHSCNDDCPGSDQDFREVQSEFSDIPRGRKFKWKTY 631
 Db 708 TVVEDODLCKLYCIABGDFPFSSLSNKKVKGDTGTPCSDSRNNVDCIGELSVSVSAHMPQPPKEDLFL 346
 Db 632 -RGGVYKACSLTCLAGNFYTERAAVVDGTPCSDPVDICVSGECKHVCGRVLSGDL 690
 Qy 768 VEDVCGVGNNSACTIHRGLYTKHHHTNOYHYMTTIPSGARSIRIYEMNVSTSVSRN 827
 Db 691 REDKCRVCGGDSACETIEGVSPAGIEDVWIIPKGSVHIFIQDLNLSHLALXG 750
 Qy 828 ALRYYLNGHWTVDMGRYKFSCTTFDYRSNNEPENLIATGPTNETLIVELLPOGRNPG 887
 Db 751 DQESLLLSGLPGTQPPLPLAGTTFQLRQGPQVQSLEALGPFNASLIYVWLARTELPA 810
 Qy 888 VAMEYSMPRLGTETKOPPAQPSYTAIVR-SECSVSCGG 925
 Db 811 LRVFNAP-IARDSLPP-----YSHYAPWTKCSAQACAG 844

RESULT 12
 US-09-842-469-2
 ; Sequence 2, Application US/09842469
 ; Publication No. US20030166999A1
 ; GENERAL INFORMATION:
 ; APPLICANT: PEIZER PRODUCT INC
 ; TITLE OF INVENTION: ADAMTS POLYPEPTIDES, NUCLEIC ACIDS ENCODING THEM, AND
 ; FILE OF INVENTION: USES THEREOF
 ; FILE REFERENCE: PC108738
 ; CURRENT APPLICATION NUMBER: US/09/842,469
 ; CURRENT FILING DATE: 2001-04-26
 ; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 1104
 ; TYPE: PRT
 ; ORGANISM: amino acid/human
 US-09-842-469-2
 Query Match 23.7%; Score 1240; DB 12; Length 1104;
 Best Local Similarity 31.3%; Pred. No. 4.5e-101;

Matches 294; Conservative 134; Mismatches 317; Indels 194; Gaps 24;
 Qy 84 PASSSTRCAAGLDGGRMDAENHRSGTNTGTENOTLHVLTOYDLVSAVEYDHRGDYV 143
 Db 3 PACOILRWALAL-GLGLMFVETHAFRSQ-----DEFLSSLESYEIAFTPRVHNGALL 54
 Qy 144 SHEIMHQRARRAVAVSVESLSLRKQPRHDPMDLKTSSSLVAPGGIVQTLGKTGTS 203
 Db 55 AFSPPPRRQRGTGATASRLFYKVASPSTHPLNLTGRSRLLAGHVSVEYVWTEGLAW 114
 Qy 204 VQTLPEPDEFYQGSLSRHSNS-----PSHGKFCESGTR 238
 Db 115 QRAARPH-CLYAGHLQQAOSTSHVAISTCGGLHGLIWADEEYLIETLHGGP--KGS-- 168
 Qy 239 TLKLCNSOKCPROS--VDFRAAQ-----CAHNSRRFRGRHYKWKPYTVQVADLC 286
 Db 169 -----RSPESGPHVYKSSLRHPLDTCAGVRDEKPKGRPWMLRTLKPPAR-- 218
 Qy 287 KLYCIABGDFPFSSLSNKKVKGDTGTPCSDSRNNVDCIGELSVSVSAHMPQPPKEDLFL 346
 Db 219 -----PLGNTERGQP-----GL----- 231
 Qy 347 PDEYKSLRHSLSLRHSHREELNVELTVVQKQWQNHGHENITTVYTLNMMVSALEK 406
 Db 232 -----KSVSR-----ERYVETLVADKMWVAYHGRDRVQYVLAIMNIAKLQ 276
 Qy 407 DGLMG-----KDG- 414
 Db 277 DSSLGTVNLVTRLLITLTDQPTLEITHHAGSLDSCFKWQKSVNHSHGNAIPENG 336
 Qy 415 TRHDHAILTGIDICSWNKEPCDTLGFAPISGKSVRSCTINEDTGLGTAFTIAHESGH 474
 Db 337 ANHDTAVLITRYDICIYKPKCTGLAPVGGCERSCSVNEDIGLATAFTIAHETGH 396
 Qy 475 NFGMTHDGEKMC---KKSEGNIMSPTLAGRNQVFSMSPCSQYLRKFLSTQAICLADQ 531
 Db 397 TFGMNDHGVNCGARGQDPFAMAAHITMKTNPFWSSCSRDYITSLDGLGLCLNRR 456
 Qy 532 PRPVEKYEPKLPGLDYDANTQCKWQFGEKAKLCLMDFKDICKALWCHRIGRCKETRF 591
 Db 457 P-PRODVPYTPAQYADABQCFQGVKRSCKYK---EVCSELWCLSKNRCITNS 512
 Qy 592 MPAAGETIC-GHDM---WCRGGQCVKYVGDGPKPTHGHWSMSSWSPCSTCGGVSHRS 647
 Db 648 RLCTNPKPSHGKFCESGTRTLKNSOKCPROSDVDFRAAQCAEHNSRRFRGRHYKWKPY 707
 Qy 572 RHCDSPRTTGGYKCLGRHRHSCNDDCPGSDQDFREVQSEFSDIPRGRKFKWKTY 631
 Db 708 TVVEDODLCKLYCIABGDFPFSSLSNKKVKGDTGTPCSDSRNNVDCIGELSVSVSAHMPQPPKEDLFL 346
 Db 632 -RGGVYKACSLTCLAGNFYTERAAVVDGTPCSDPVDICVSGECKHVCGRVLSGDL 690
 Qy 768 VEDVCGVGNNSACTIHRGLYTKHHHTNOYHYMTTIPSGARSIRIYEMNVSTSVSRN 827
 Db 691 REDKCRVCGGDSACETIEGVSPAGIEDVWIIPKGSVHIFIQDLNLSHLALXG 750
 Qy 828 ALRYYLNGHWTVDMGRYKFSCTTFDYRSNNEPENLIATGPTNETLIVELLPOGRNPG 887
 Db 751 DQESLLLSGLPGTQPPLPLAGTTFQLRQGPQVQSLEALGPFNASLIYVWLARTELPA 810
 Qy 888 VAMEYSMPRLGTETKOPPAQPSYTAIVR-SECSVSCGG 925
 Db 811 LRVFNAP-IARDSLPP-----YSHYAPWTKCSAQACAG 844

RESULT 13
 US-09-858-068-2
 ; Sequence 2, Application US/09858068
 ; Patent No. US20020076778A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kapeller-Liebertmann, Rosana

```

; APPLICANT: Cook, William James
; APPLICANT: Siles-Santiago, Immaculada
; TITLE OF INVENTION: 33428, A NOVEL HUMAN METALLOPROTEASE
; TITLE OF INVENTION: FAMILY MEMBER AND USES THEREOF
; FILE REFERENCE: 10448-057001
; CURRENT APPLICATION NUMBER: US/09/858, 068
; CURRENT FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: 60/204,159
; PRIOR FILING DATE: 2000-05-15
; PRIOR FILING DATE: 2000-05-15
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1133
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-858-068-2

Query Match      23.6%; Score 1234; DB 9; Length 1133;
Best Local Similarity 30.3%; Pred. No. 1.6e-100;
Matches 310; Conservative 138; Mismatches 332; Indels 244; Gaps 29;

QY 24 QVSPGRSHORNGRSQQLASPPRLISGPRRLTAMSPLSAGCTVHGTRSSAWEPER 83
DB 31 QKEFGDAKAVNRGGCTVGAAG-----SG-----NGETC---GHVSMK----- 67
QY 84 PASSSTGAGLQDGRMDPAGNHSQOVTGTENQTLVLTQYDLVSAVEYDHRGDYV 143
DB 68 PACQILRWALAL-GLGLAFVETHAPRSQ-----DEFLSLESYEIAFPTRVDHNGALL 119
QY 144 SHEIMHORRRRAVAVESVLSHLRKGPRHDFHMDLTSSSVIAPGFIVOTLGTGTS 203
DB 120 AFSPPPRQRGRGTGATAESLFYKVASPSTHFLNLTSSRLLAGHVSVEYWTREGLAW 179
QY 204 VQTLPPEDFCFYQGSLSRSHNS-----PSHGGRFCEGSTR 238
DB 180 QRAARPH--CLYAGHLQOQASSSSHVAISTCGUGHLGLI VADEEYLIETLHGPP--KGS-- 233
QY 239 TLKLCNSQKCPDSD--VDFAAQ-----CAHNSRRPRGRHYKWPYQTOVEADLC 286
DB 234 -----RSPESGPHVYKRSLSRHPHLDTAGVRDEKPKWGRPWLRTLPKPPAR-- 283
QY 287 KLYCIAEGDFEFLSNKVKDGTCPSEDNRNVCIDGICELSVSVSTSAHMPQPKEDLFI 346
DB 284 -----PLCNETERGOF-----GL----- 296
QY 347 PDEYKSLRHKRSLLSHRNEELNETLVVVDKMKMQNHGHEINITTYVLTILMWSALFX 406
DB 297 -----KRSVSR-----ERVETLVADKMWAVHGRDVEQYVLAIMVIAKLFQ 341
QY 407 DGLMG-----KDG-- 414
DB 342 DSSIGSTVMIILVTLILLTDOPTLITHAGKSLDSFCWKOKSIVNHSGHGNAIPENG 401
QY 415 TRHDHAILTLGLDLSWNRNPECDTLGAPISGMSKYRSCYINEDTGLGLAFTIAHESGH 474
DB 402 ANHDTAVLITRYDICTYKPKCGTGLAPGVGNCERSCSVNEDIGLATFTIAHEIGH 461
QY 475 NFGMHDGEGNCKSEGNIMSPSLAGRNQVFGSPCRQVLRHLFLSTAQAICLADQPKP 534
DB 462 TFGNHDHGVGNSC--GAPKJMAHMTKNTNFWSSCRDYITPSFLOGLGLCLNRP-P 518
QY 535 VKEYTYPEKLGELDYANTQKQWQGEKAKLMDPKDICKALMCHRIKCKETKMPA 594
DB 519 RQDEVPTVAVGQYADQPCQRPQGVASRCKYV---EVSELMCLSKSNCRNTNSIPA 575
QY 595 AEGTIC-GHNM---WCRGQCVKYVDEGPKPTHGHSWNSWSPSCRTCGGVSHRSLC 650
DB 576 AEGTLQTHITDKGWCYKRVVCPGSR--PEGVDGAWGTWPDGCSRTCGGVSSSSRHC 634
QY 651 INPSPSHGKVCBESRTTLKLCNSQKCPDSDVDFRAAOCAHNSRRFRGRHYKWPYQV 710
DB 651 INPSPSHGKVCBESRTTLKLCNSQKCPDSDVDFRAAOCAHNSRRFRGRHYKWPYQV 710

```

```

635 DSPRPTIGKYCLGERRRHRSNTDDCPGSGQDFREVQCSEFDSIPFRGKYKWKTY--RG 693
711 EDQDLCKLYCIAEGDFEFLSNKVKDGTCPSEDNRNVCIDGICERVCNGLVDAVED 770
694 GGKACSLTCLASGFNFYTERAAAVDGTFCREDTVDICVSGCKHVCGRVIGSLDRE 753
771 VCGVQNGNSACTIHRGLYTKHHHTNOYVHWVTIPSGARSIRIYEMNVSTYSVRLNR 830
754 KCRVCGGDSACSTIEGVFSPASPGAGYEDVWVWPKGVSHIFQDLNLSLSHLKAL 813
831 RYVLNGHMTVDWFCRYKFSGTTDFDYRRSYNEPENLIATGPTNETLIV----- 877
814 SLLLEGULPTQPRLPLAGTTTFLQFGQGPQVQSLEALGPINASLVNVTGRUGARCS 873
878 -----ELLFOQNFVAVMEYSMPRLGTSKOPPAQPSYTAIVR--SECSVS 921
874 NRLHATLPLAAAPFLOVLARTELPAIRYFNAP--IARDSLPP-----YSMHYAPWKCSAQ 928
922 CCGG 925
929 CAGG 932

RESULT 14
US-09-858-081-2
; Sequence 2, Application US/09858081
; Patent No. US20020072490A1
; GENERAL INFORMATION:
; APPLICANT: Kapellier-Libermann, Rosana
; APPLICANT: Siles-Santiago, Immaculada
; TITLE OF INVENTION: 33428, A NOVEL HUMAN METALLOPROTEASE
; TITLE OF INVENTION: FAMILY MEMBER AND USES THEREOF
; FILE REFERENCE: 10448-057001
; CURRENT APPLICATION NUMBER: US/09/858, 081
; CURRENT FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: 60/204,159
; PRIOR FILING DATE: 2000-05-15
; PRIOR APPLICATION NUMBER: 60/204,160
; PRIOR FILING DATE: 2000-05-15
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1068
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-858-081-2

Query Match      23.3%; Score 1222.5; DB 9; Length 1068;
Best Local Similarity 30.6%; Pred. No. 1.6e-99;
Matches 295; Conservative 133; Mismatches 315; Indels 221; Gaps 25;

QY 84 PASSSTGAGLQDGRMDPAGNHSQOVTGTENQTLVLTQYDLVSAVEYDHRGDYV 143
DB 3 PACQILRWALAL-GLGLAFVETHAPRSQ-----DEFLSLESYEIAFPTRVDHNGALL 54
QY 144 SHEIMHORRRRAVAVESVLSHLRKGPRHDFHMDLTSSSVIAPGFIVOTLGTGTS 203
DB 55 AFSPPPRQRGRGTGATAESLFYKVASPSTHFLNLTSSRLLAGHVSVEYWTREGLAW 114
QY 204 VQTLPPEDFCFYQGSLSRSHNS-----PSHGGRFCEGSTR 238
DB 115 QRAARPH--CLYAGHLQOQASSSSHVAISTCGUGHLGLI VADEEYLIETLHGPP--KGS-- 168
QY 239 TLKLCNSQKCPDSD--VDFAAQ-----CAHNSRRPRGRHYKWPYQTOVEADLC 286
DB 169 -----RSPESGPHVYKRSLSRHPHLDTAGVRDEKPKWGRPWLRTLPKPPAR-- 218
QY 287 KLYCIAEGDFEFLSNKVKDGTCPSEDNRNVCIDGICELSVSVSTSAHMPQPKEDLFI 346
DB 219 -----PLCNETERGOF-----GL----- 231
QY 347 PDEYKSLRHKRSLLSHRNEELNETLVVVDKMKMQNHGHEINITTYVLTILMWSALFX 406

```

